

# Action Theory Evolution

Ivan José Varzinczak  
Meraka Institute  
CSIR Pretoria, South Africa  
`ivan.varzinczak@meraka.org.za`

## Abstract

Like any other logical theory, domain descriptions in reasoning about actions may evolve, and thus need revision methods to adequately accommodate new information about the behavior of actions. The present work is about changing action domain descriptions in propositional dynamic logic. Its contribution is threefold: first we revisit the semantics of action theory contraction that has been done in previous work, giving more robust operators that express minimal change based on a notion of distance between Kripke-models. Second we give algorithms for syntactical action theory contraction and establish their correctness w.r.t. our semantics. Finally we state postulates for action theory contraction and assess the behavior of our operators w.r.t. them. Moreover, we also address the revision counterpart of action theory change, showing that it benefits from our semantics for contraction.

# Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
<b>2</b>	<b>Logical Preliminaries</b>	<b>6</b>
2.1	Action Theories in Dynamic Logic . . . . .	6
2.2	Essential Atoms . . . . .	11
2.3	Prime Valuations . . . . .	12
2.4	Closeness between Models . . . . .	13
<b>3</b>	<b>Semantics of Action Theory Change</b>	<b>13</b>
3.1	Model Contraction of Executability Laws . . . . .	14
3.2	Model Contraction of Effect Laws . . . . .	16
3.3	Model Contraction of Static Laws . . . . .	20
<b>4</b>	<b>Syntactic Operators for Contraction of Laws</b>	<b>21</b>
4.1	Contracting Executability Laws . . . . .	22
4.2	Contracting Effect Laws . . . . .	23
4.3	Contracting Static Laws . . . . .	27
<b>5</b>	<b>Correctness of the Operators</b>	<b>28</b>
5.1	Two Counter-Examples . . . . .	28
5.2	Modular Theories . . . . .	30
5.3	Correctness Under Modularity . . . . .	31
<b>6</b>	<b>Assessment of Postulates for Change</b>	<b>32</b>
6.1	Contraction or Erasure? . . . . .	32
6.2	The Postulates . . . . .	33
<b>7</b>	<b>A Semantics for Action Theory Revision</b>	<b>37</b>
7.1	Revising a Model by a Static Law . . . . .	38
7.2	Revising a Model by an Effect Law . . . . .	40
7.3	Revising a Model by an Executability Law . . . . .	41
7.4	Revising Sets of Models . . . . .	42
<b>8</b>	<b>Related Work</b>	<b>43</b>
<b>9</b>	<b>Comments</b>	<b>46</b>
9.1	Other Distance Notions . . . . .	46
9.2	Inducing Executability . . . . .	47

<b>10 Concluding Remarks</b>	<b>49</b>
<b>A Proof of Theorem 5.2</b>	<b>56</b>
<b>B Proof of Theorem 5.3</b>	<b>61</b>

# 1 Introduction

Consider an intelligent agent designed to perform rationally in a dynamic world, and suppose she should reason about the dynamics of an automatic coffee machine (Figure 1). Suppose, for example, that the agent believes that coffee is always a hot beverage. Suppose now that some day she gets a coffee and observes that it is cold. In such a case, the agent must change her beliefs about the relation between the propositions “I hold a coffee” and “I hold a hot beverage”. This example is an instance of the problem of changing propositional belief bases and is largely addressed in the literature about belief change [15] and belief update [31].

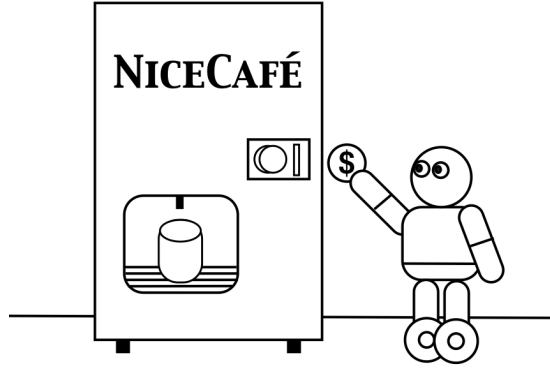


Figure 1: The coffee deliverer agent.

Next, let our agent believe that whenever she buys a coffee from the machine, she gets a hot beverage. This means that in every state of the world that follows the execution of buying a coffee, the agent possesses a hot beverage. Then, in a situation where the machine is running out of cups, after buying, the coffee runs through the shelf and the agent does not hold a hot beverage in her hands.

Imagine now that the agent never considered any relation between buying a coffee on the machine and its service availability, in the sense that the agent always believed that buying does not prevent other users from using the machine. Nevertheless, someday our agent is queuing to buy a coffee and observes that just after the agent before her has bought, the machine went out of order (maybe due to a lack of coffee powder).

Completing our agent’s struggle in discovering the intricacies of a coffee machine, suppose she always believed that if she has a token, then it is

possible to buy coffee, provided that some other conditions like being close enough to the button, having a free hand, etc, are satisfied. However, during a blackout, the agent, even with a token, does not manage to buy her coffee.

The last three examples illustrate situations where changing the beliefs about the *behavior* of the action of buying coffee is mandatory. In the first one, buying coffee, once believed to be deterministic, has now to be seen as nondeterministic, or alternatively to have a different outcome in a more specific context (e.g. if there is no cup in the machine). In the second example, buying a coffee is now known to have side-effects (ramifications) one was not aware of. Finally, in the last example, the executability of the action under concern is questioned in the light of new information showing a context that was not known to preclude its execution.

Such cases of theory change are very important when one deals with logical descriptions of dynamic domains: it may always happen that one discovers that an action actually has a behavior that is different from that one has always believed it had.

Up to now, theory change has been studied mainly for knowledge bases in classical logics, both in terms of revision and update. Since the work by Fuhrmann [14], only in a few recent studies has it been considered in the realm of modal logics, viz. in epistemic logic [19] and in dynamic logics [21]. Recently some studies have investigated revision of beliefs about *facts* of the world [47, 28] or the agent's goals [46]. In our scenario, this would concern for instance the truth of *token* in a given state: the agent believes that she has a token, but is actually wrong about that. Then she might subsequently be forced to revise her beliefs about the current state of affairs or change her goals according to what she can perform in that state. Such belief revision operations do not modify the agent's beliefs about the *action laws*. In opposition to that, here we are interested exactly in such modifications. Starting with Baral and Lobo's work [4], some recent studies have been done on that issue [12, 13] for domain descriptions in action languages [16].

We here take a step further in this direction and propose a method based on that given by Herzig *et al.* [21] that is more robust by integrating a notion of *minimal change* and complying with *postulates* of theory change.

The present text is structured as follows: in Section 2 we establish the formal background that will be used throughout this work. Sections 3–5 are the core of the work: in Section 3 we present the central definitions for a semantics of action theory change, Section 4 is devoted to its syntactical counterpart while Section 5 to the proof of its correspondence with the semantics. In Section 6 we discuss some postulates for contraction/erasure and

then present a semantics for action theory revision (Section 7). In Section 8 we address existing work in the field. After making some comments on our method (Section 9), we finish with some conclusions and future directions of research.

## 2 Logical Preliminaries

Following the tradition in the reasoning about actions (RAA) community, we consider action theories to be finite collections of statements that have the particular form:

- if *context*, then *effect* after *every execution* of *action* (effect laws);
- if *precondition*, then *action executable* (executability laws).

Statements mentioning no action at all represent laws about the underlying structure of the world, i.e., its possible states (static laws).

Several logical frameworks have been proposed to formalize such statements. Among the most prominent ones are the Situation Calculus [39, 45], the family of Action Languages [16, 30, 17], the Fluent Calculus [49, 50], and the dynamic logic-based approaches [10, 6, 57]. Here we opt to formalize action theories using a version of Propositional Dynamic Logic (PDL) [20].

### 2.1 Action Theories in Dynamic Logic

Let  $\mathcal{Act} = \{a_1, a_2, \dots\}$  be the set of all *atomic action constants* of a given domain. An example of atomic action is *buy*. To each atomic action  $a$  there is associated a modal operator  $[a]$ .<sup>1</sup>

$\mathcal{Prop} = \{p_1, p_2, \dots\}$  denotes the set of all *propositional constants*, also called *fluents* or *atoms*. Examples of those are *token* (“the agent has a token”) and *coffee* (“the agent holds a coffee”). The set of all literals is  $\mathcal{Lit} = \{\ell_1, \ell_2, \dots\}$ , where each  $\ell_i$  is either  $p$  or  $\neg p$ , for some  $p \in \mathcal{Prop}$ . If  $\ell = \neg p$ , then we identify  $\neg \ell$  with  $p$ . By  $|\ell|$  we denote the atom in  $\ell$ .

We use small Greek letters  $\varphi, \psi, \dots$  to denote *Boolean formulas*. They are recursively defined in the usual way:

$$\varphi ::= p \mid \top \mid \perp \mid \neg \varphi \mid \varphi \wedge \varphi \mid \varphi \vee \varphi \mid \varphi \rightarrow \varphi \mid \varphi \leftrightarrow \varphi$$

---

<sup>1</sup>We here suppose that our multimodal logic is independently axiomatized [32], i.e., the logic is a fusion and there is no interaction between the modal operators. This is a requirement to achieve modularity of action theories [25] (see further).

$\mathfrak{Fml}$  is the set of all Boolean formulas. An example of a Boolean formula is  $coffee \rightarrow hot$ . A propositional valuation  $v$  is a *maximally consistent* set of literals. We denote by  $v \models \varphi$  the fact that  $v$  satisfies a propositional formula  $\varphi$ . By  $val(\varphi)$  we denote the set of all valuations satisfying  $\varphi$ .  $\models_{\text{CPL}}$  denotes the classical consequence relation.  $Cn(\varphi)$  denotes all logical consequences of  $\varphi$  in classical propositional logic.

If  $\varphi$  is a propositional formula,  $atm(\varphi)$  denotes the set of elementary atoms *actually* occurring in  $\varphi$ . For example,  $atm(\neg p_1 \wedge (\neg p_1 \vee p_2)) = \{p_1, p_2\}$ .

For  $\varphi$  a Boolean formula,  $IP(\varphi)$  denotes the set of its *prime implicants* [43], i.e., the weakest terms (conjunctions of literals) that imply  $\varphi$ . As an example,  $IP(p_1 \oplus p_2) = \{p_1 \wedge \neg p_2, \neg p_1 \wedge p_2\}$ . For more on prime implicants, their properties and how to compute them see the chapter by Marquis [36]. By  $\pi$  we denote a prime implicant, and given  $\ell$  and  $\pi$ ,  $\ell \in \pi$  abbreviates ‘ $\ell$  is a literal of  $\pi$ ’.

We denote complex formulas (possibly with modal operators) by  $\Phi, \Psi, \dots$ . They are recursively defined in the following way:

$$\Phi ::= \varphi \mid [a]\Phi \mid \neg\Phi \mid \Phi \wedge \Phi \mid \Phi \vee \Phi \mid \Phi \rightarrow \Phi \mid \Phi \leftrightarrow \Phi$$

$\langle a \rangle$  is the dual operator of  $[a]$ , defined as  $\langle a \rangle\Phi =_{\text{def}} \neg[a]\neg\Phi$ . An example of a complex formula is  $\neg coffee \rightarrow [buy]coffee$ .

The semantics is that of PDL without the  $*$  operator, which amounts to multimodal logic  $K_n$  [42]. In the following we will refer to PDL but our underlying logical formalism is essentially the simpler multimodal logic  $K_n$ , which turns out to be expressive enough for our purposes here.

**Definition 2.1 (PDL-model)** A PDL-model is a tuple  $\mathcal{M} = \langle W, R \rangle$  where  $W$  is a set of valuations (also called possible worlds), and  $R$  maps action constants  $a$  to accessibility relations  $R_a \subseteq W \times W$ .

As an example, for  $\mathfrak{Act} = \{a_1, a_2\}$  and  $\mathfrak{Prop} = \{p_1, p_2\}$ , we have the PDL-model  $\mathcal{M} = \langle W, R \rangle$ , where

$$\begin{aligned} W &= \{\{p_1, p_2\}, \{p_1, \neg p_2\}, \{\neg p_1, p_2\}\}, \\ R(a_1) &= \left\{ \begin{array}{l} (\{p_1, p_2\}, \{p_1, \neg p_2\}), (\{p_1, p_2\}, \{\neg p_1, p_2\}), \\ (\{p_1, \neg p_2\}, \{p_1, \neg p_2\}), (\{p_1, \neg p_2\}, \{\neg p_1, p_2\}) \end{array} \right\} \\ R(a_2) &= \{(\{p_1, p_2\}, \{\neg p_1, p_2\}), (\{\neg p_1, p_2\}, \{\neg p_1, p_2\})\} \end{aligned}$$

Figure 2 gives a graphical representation of  $\mathcal{M}$ .<sup>2</sup>

---

<sup>2</sup>Notice that our notion of PDL-model does not follow the standard notion from modal

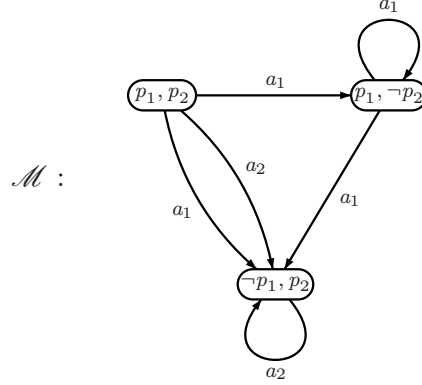


Figure 2: Example of a PDL-model for  $\mathfrak{Act} = \{a_1, a_2\}$ , and  $\mathfrak{Prop} = \{p_1, p_2\}$ .

**Definition 2.2 (Truth conditions)** Given a PDL-model  $\mathcal{M} = \langle W, R \rangle$ ,

- $\models_w^{\mathcal{M}} p$  ( $p$  is true at world  $w$  of model  $\mathcal{M}$ ) if  $w \Vdash p$  (the valuation  $w$  satisfies  $p$ , i.e.,  $p \in w$ );
- $\models_w^{\mathcal{M}} [a]\Phi$  if  $\models_{w'}^{\mathcal{M}} \Phi$  for every  $w'$  s.t.  $(w, w') \in R_a$ ;
- $\models_w^{\mathcal{M}} \Phi \wedge \Psi$  if  $\models_w^{\mathcal{M}} \Phi$  and  $\models_w^{\mathcal{M}} \Psi$ ;
- $\models_w^{\mathcal{M}} \Phi \vee \Psi$  if  $\models_w^{\mathcal{M}} \Phi$  or  $\models_w^{\mathcal{M}} \Psi$ , or both;
- $\models_w^{\mathcal{M}} \neg\Phi$  if  $\not\models_w^{\mathcal{M}} \Phi$ , i.e., not  $\models_w^{\mathcal{M}} \Phi$ ;
- truth conditions for the other connectives are as usual.

By  $\mathcal{M}$  we will denote a set of PDL-models.

A PDL-model  $\mathcal{M}$  is a model of  $\Phi$  (denoted  $\models^{\mathcal{M}} \Phi$ ) if and only if for all  $w \in W$ ,  $\models_w^{\mathcal{M}} \Phi$ . In the model depicted in Figure 2, we have  $\models^{\mathcal{M}} p_1 \rightarrow [a_2]p_2$  and  $\models^{\mathcal{M}} p_1 \vee p_2$ .

**Definition 2.3 (Global consequence)**  $\mathcal{M}$  is a model of a set of formulas  $\Sigma$  (noted  $\models^{\mathcal{M}} \Sigma$ ) if and only if  $\models^{\mathcal{M}} \Phi$  for every  $\Phi \in \Sigma$ . A formula  $\Phi$  is a consequence of a set of global axioms  $\Sigma$  in the class of all PDL-models (noted  $\Sigma \models_{\text{PDL}} \Phi$ ) if and only if for every PDL-model  $\mathcal{M}$ , if  $\models^{\mathcal{M}} \Sigma$ , then  $\models^{\mathcal{M}} \Phi$ .

---

logics: here no two worlds satisfy the same valuation. This is a pragmatic choice (see Section 4). Nevertheless, all we are about to state in the sequel can be straightforwardly formulated for standard PDL models as well.



With PDL we can state laws describing the behavior of actions. One way of doing this is by stating some formulas as global axioms.<sup>3</sup> As usually done in the RAA community, we here distinguish three types of laws. The first kind of statements are *static laws*, which are Boolean formulas that must hold in every possible state of the world.

**Definition 2.4 (Static Law)** *A static law is a formula  $\varphi \in \mathfrak{Sml}$ .*

An example of a static law is  $\text{coffee} \rightarrow \text{hot}$ , saying that if the agent holds a coffee, then she holds a hot beverage. The set of all static laws of a domain is denoted by  $\mathcal{S} \subseteq \mathfrak{Sml}$ . In our example we will have  $\mathcal{S} = \{\text{coffee} \rightarrow \text{hot}\}$ .

The second kind of action law we consider is given by the *effect laws*. These are formulas relating an action to its effects, which can be conditional.

**Definition 2.5 (Effect Law)** *An effect law for action  $a$  is of the form  $\varphi \rightarrow [a]\psi$ , where  $\varphi, \psi \in \mathfrak{Sml}$ .*

The consequent  $\psi$  is the effect which always obtains when action  $a$  is executed in a state where the antecedent  $\varphi$  holds. If  $a$  is a nondeterministic action, then the consequent  $\psi$  is typically a disjunction. An example of an effect law is  $\neg \text{coffee} \rightarrow [\text{buy}]\text{coffee}$ , saying that in a situation where the agent has no coffee, after buying, the agent has a coffee. If  $\psi$  is inconsistent, then we have a special kind of effect law that we call an *inexecutability law*. For example, we could also have  $\neg \text{token} \rightarrow [\text{buy}]\perp$ , expressing that *buy* cannot be executed if the agent has no token.

The set of effect laws of a domain is denoted by  $\mathcal{E}$ . In our coffee machine scenario, we could have for example:

$$\mathcal{E} = \{ \neg \text{coffee} \rightarrow [\text{buy}]\text{coffee}, \text{token} \rightarrow [\text{buy}]\neg \text{token}, \neg \text{token} \rightarrow [\text{buy}]\perp \}$$

Finally, we also define *executability laws*, which stipulate the context where an action is guaranteed to be executable. In PDL, the operator  $\langle a \rangle$  is used to express executability.  $\langle a \rangle \top$  thus reads “the execution of  $a$  is possible”.

**Definition 2.6 (Executability Law)** *An executability law for action  $a$  is of the form  $\varphi \rightarrow \langle a \rangle \top$ , where  $\varphi \in \mathfrak{Sml}$ .*

---

<sup>3</sup>An alternative to that is given by Castilho *et al.* [6], with laws being stated with the aid of an extra universal modality and local consequence being thus considered.

For instance,  $token \rightarrow \langle buy \rangle \top$  says that buying can be executed whenever the agent has a token. The set of all executability laws of a given domain is denoted by  $\mathcal{X}$ . In our scenario example we would have  $\mathcal{X} = \{token \rightarrow \langle buy \rangle \top\}$ .

With our three basic types of laws, we are able to define action theories:

**Definition 2.7 (Action Theory)** *Given a domain and any (possibly empty) sets of laws  $\mathcal{S}$ ,  $\mathcal{E}$ , and  $\mathcal{X}$ ,  $\mathcal{T} = \mathcal{S} \cup \mathcal{E} \cup \mathcal{X}$  is an action theory.*

For given action  $a$ ,  $\mathcal{E}_a$  (resp.  $\mathcal{X}_a$ ) will denote the set of only those effect (resp. executability) laws about  $a$ .  $\mathcal{T}_a = \mathcal{S} \cup \mathcal{E}_a \cup \mathcal{X}_a$  is then the action theory for  $a$ .<sup>4</sup>

For the sake of clarity, we abstract here from the frame and ramification problems, and suppose the agent's theory already entails all the relevant frame axioms. We could have used any suitable solution to the frame problem, like e.g. the dependence relation [6], which is used in the work of Herzig *et al.* [21], or a kind of successor state axioms in a slightly modified setting [11]. To make the presentation more clear to the reader, here we do not bother with a solution to the frame problem and just assume all frame axioms can be inferred from the theory. Actually we can suppose that all intended frame axioms are automatically recovered and stated in the theory, more specific, in the set of effect laws.<sup>5</sup> Hence the action theory of our example will be:

$$\mathcal{T} = \left\{ \begin{array}{l} coffee \rightarrow hot, token \rightarrow \langle buy \rangle \top, \\ \neg coffee \rightarrow [buy] coffee, \\ token \rightarrow [buy] \neg token, \neg token \rightarrow [buy] \perp, \\ coffee \rightarrow [buy] coffee, hot \rightarrow [buy] hot \end{array} \right\}$$

(We have not stated the frame axiom  $\neg token \rightarrow [buy] \neg token$  because it can be trivially deduced from the inexecutability law  $\neg token \rightarrow [buy] \perp$ .)

Figure 3 below shows a PDL-model for the theory  $\mathcal{T}$ .

Given an action theory  $\mathcal{T}$ , sometimes it will be useful to consider models whose possible worlds are *all* the possible worlds allowed by  $\mathcal{T}$ :

---

<sup>4</sup>Notice that for  $a_1, a_2 \in \mathfrak{Act}$ ,  $a_1 \neq a_2$ , the intuition is indeed that  $\mathcal{T}_{a_1}$  and  $\mathcal{T}_{a_2}$  overlap only on  $\mathcal{S}$ , i.e., the only laws that are common to both  $\mathcal{T}_{a_1}$  and  $\mathcal{T}_{a_2}$  are the laws about the structure of the world. This requirement is somehow related with the logic being independently axiomatized (see above).

<sup>5</sup>Frame axioms are a special type of effect law, having the form  $\ell \rightarrow [a]\ell$ , for  $\ell \in \mathfrak{Lit}$ .

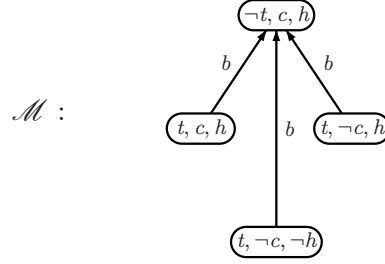


Figure 3: A model for our coffee machine scenario:  $b$ ,  $t$ ,  $c$ , and  $h$  stand for, respectively, *buy*, *token*, *coffee*, and *hot*.

**Definition 2.8 (Big Model)** Let  $\mathcal{T} = \mathcal{S} \cup \mathcal{E} \cup \mathcal{X}$  be an action theory.  $\mathcal{M}_{big} = \langle W_{big}, R_{big} \rangle$  is the big model of  $\mathcal{T}$  if and only if:

- $W_{big} = \text{val}(\mathcal{S})$ ; and
- $R_{big} = \bigcup_{a \in \mathcal{A}_{ct}} R_a$  s.t.  $R_a = \{(w, w') : \text{for all } \varphi \rightarrow [a]\psi \in \mathcal{E}_a, \text{ if } \models_w^{\mathcal{M}} \varphi, \text{ then } \models_{w'}^{\mathcal{M}} \psi\}$ .

Figure 4 below shows the big model of  $\mathcal{T}$ .

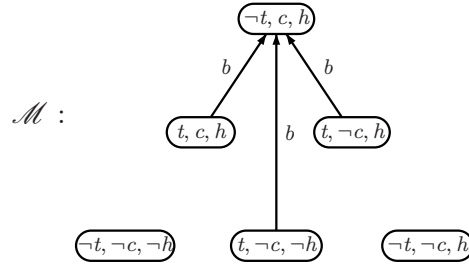


Figure 4: The big model for the coffee machine scenario.

## 2.2 Essential Atoms

An atom  $p$  is *essential* to a formula  $\varphi$  if and only if  $p \in \text{atm}(\varphi')$  for every  $\varphi'$  such that  $\models_{\text{CPL}} \varphi \leftrightarrow \varphi'$ . For instance,  $p_1$  is essential to  $\neg p_1 \wedge (\neg p_1 \vee p_2)$ . Given  $\varphi$ ,  $\text{atm}!(\varphi)$  denotes the set of essential atoms of  $\varphi$ . (If  $\varphi$  is not contingent, i.e.,  $\varphi$  is a tautology or a contradiction, then  $\text{atm}!(\varphi) = \emptyset$ .)

Given  $\varphi$  a Boolean formula,  $\varphi^*$  is the set of all formulas  $\varphi'$  such that  $\varphi \models_{\text{CPL}} \varphi'$  and  $\text{atm}(\varphi') \subseteq \text{atm}(\varphi)$ . For instance,  $p_1 \vee p_2 \notin p_1^*$ , as  $p_1 \models_{\text{CPL}} p_1 \vee p_2$  but  $\text{atm}(p_1 \vee p_2) \not\subseteq \text{atm}(p_1)$ . Clearly,  $\text{atm}(\bigwedge \varphi^*) = \text{atm}(\bigwedge \varphi^*)$ , moreover whenever  $\models_{\text{CPL}} \varphi \leftrightarrow \varphi'$  is the case, then  $\text{atm}(\varphi) = \text{atm}(\varphi')$  and also  $\varphi^* = \varphi'^*$ .

**Theorem 2.1 (Least atom-set theorem [41])** *Given  $\varphi$  a propositional formula,  $\models_{\text{CPL}} \varphi \leftrightarrow \bigwedge \varphi^*$ , and for every  $\varphi'$  s.t.  $\models_{\text{CPL}} \varphi \leftrightarrow \varphi'$ ,  $\text{atm}(\varphi^*) \subseteq \text{atm}(\varphi')$ .*

A proof of this theorem is given by Makinson [35] and we do not state it here. Essentially, the theorem establishes that for every formula  $\varphi$ , there is a unique least set of elementary atoms such that  $\varphi$  may equivalently be expressed using only letters from that set.<sup>6</sup> Hence,  $\text{Cn}(\varphi) = \text{Cn}(\varphi^*)$ .

### 2.3 Prime Valuations

Given a valuation  $v$ ,  $v' \subseteq v$  is a *subvaluation*. Given a set of valuations  $W$ , a subvaluation  $v'$  *satisfies* a propositional formula  $\varphi$  modulo  $W$  (noted  $v' \models_W \varphi$ ) if and only if  $v \models \varphi$  for all  $v \in W$  such that  $v' \subseteq v$ .

We say that a subvaluation  $v$  *essentially satisfies*  $\varphi$  (modulo  $W$ ), noted  $v \models_W^! \varphi$ , if and only if  $v \models_W \varphi$  and  $\{|\ell| : \ell \in v\} \subseteq \text{atm}(\varphi)$ . If  $v \models_W^! \varphi$ , we call  $v$  an *essential subvaluation* of  $\varphi$  (modulo  $W$ ).

**Definition 2.9 (Prime Subvaluation)** *Let  $\varphi$  be a propositional formula and  $W$  a set of valuations. A subvaluation  $v$  is a prime subvaluation of  $\varphi$  (modulo  $W$ ) if and only if  $v \models_W^! \varphi$  and there is no  $v' \subseteq v$  s.t.  $v' \models_W^! \varphi$ .*

Our notion of prime subvaluation is closely related to Veltman's definition of *basis* for a formula [54].<sup>7</sup> A prime subvaluation of a formula  $\varphi$  is thus one of the weakest states of truth in which  $\varphi$  is true. Hence, prime subvaluations are just another way of seeing prime implicants [43] of  $\varphi$ . By  $\text{base}(\varphi, W)$  we will denote the set of all prime subvaluations of  $\varphi$  modulo  $W$ .

**Theorem 2.2** *Let  $\varphi \in \mathfrak{Fml}$  and  $W$  be a set of valuations. Then for all  $w \in W$ ,  $w \models \varphi$  if and only if  $w \models \bigvee_{v \in \text{base}(\varphi, W)} \bigwedge_{\ell \in v} \ell$ .*

<sup>6</sup>The dual notion, i.e., that of redundant atoms is also addressed in the literature [22], with similar purposes.

<sup>7</sup>The author is indebted to Andreas Herzig for pointing this out.

**Proof:** Right to left direction is straightforward. For the left to right direction, if  $w \Vdash \varphi$ , then  $w \Vdash \varphi^*$ . Let  $w' \subseteq w$  be the least subset of  $w$  still satisfying  $\varphi^*$ . Clearly,  $w'$  is a prime subvaluation of  $\varphi$  modulo  $W$ , and then because  $w \Vdash \bigwedge_{\ell \in w'} \ell$ , the result follows. ■

## 2.4 Closeness between Models

When contracting a formula from a model, we will perform a change in its structure. Because there can be several different ways of modifying a model (not all of them minimal), we need a notion of distance between models to identify those that are closest to the original one.

As we are going to see in more depth in what follows, changing a model amounts to modifying its possible worlds or its accessibility relation. Hence, the distance between two PDL-models will depend upon the distance between their sets of worlds and accessibility relations. These here will be based on the *symmetric difference* between sets, defined as  $X \dot{-} Y = (X \setminus Y) \cup (Y \setminus X)$ .

**Definition 2.10 (Closeness between PDL-Models)** Let  $\mathcal{M} = \langle W, R \rangle$  be a model. Then  $\mathcal{M}' = \langle W', R' \rangle$  is at least as close to  $\mathcal{M}$  as  $\mathcal{M}'' = \langle W'', R'' \rangle$ , noted  $\mathcal{M}' \preceq_{\mathcal{M}} \mathcal{M}''$ , if and only if

- either  $W \dot{-} W' \subseteq W \dot{-} W''$
- or  $W \dot{-} W' = W \dot{-} W''$  and  $R \dot{-} R' \subseteq R \dot{-} R''$

Although simple, this notion of closeness is sufficient for our purposes here, as we will see in the sequel. Notice that other distance notions could have been considered as well, like e.g. the *cardinality* of symmetric differences. (See Section 9 for a discussion on this.)

## 3 Semantics of Action Theory Change

When admitting the possibility of a law  $\Phi$  failing, one must ensure that  $\Phi$  becomes invalid, i.e., not true in at least one model of the dynamic domain. Because there can be lots of such models, we may have a *set*  $\mathcal{M}$  of models in which  $\Phi$  is (potentially) valid. Thus contracting  $\Phi$  amounts to making it no longer valid in this set of models. What are the operations that must be carried out to achieve that? Throwing models out of  $\mathcal{M}$  does not work, since  $\Phi$  will keep on being valid in all models of the remaining set. Thus one should *add* new models to  $\mathcal{M}$ . Which models? Well, models in which  $\Phi$  is

not true. But not any of such models: taking models falsifying  $\Phi$  that are too different from our original models will certainly violate minimal change.

Hence, we shall take some model  $\mathcal{M} \in \mathcal{M}$  as basis and manipulate it to get a new model  $\mathcal{M}'$  in which  $\Phi$  is not true. In dynamic logic, the removal of a law  $\Phi$  from a model  $\mathcal{M} = \langle W, R \rangle$  means modifying the possible worlds or the accessibility relation in  $\mathcal{M}$  so that  $\Phi$  becomes false. Such an operation gives as result a set  $\mathcal{M}_{\Phi}^{-}$  of models each of which is no longer a model of  $\Phi$ . But if there are several candidates, which ones should we choose? We shall take those models that are *minimal* modifications of the original  $\mathcal{M}$ , i.e., those minimal w.r.t.  $\preceq_{\mathcal{M}}$ . Note that there can be more than one  $\mathcal{M}'$  that is minimal. Hence, because adding just one of these new models is enough to invalidate  $\Phi$ , we take all possible combinations  $\mathcal{M} \cup \{\mathcal{M}'\}$  of expanding our original set of models by one of these minimal models. The result will be a *set of sets of models*. In each set of models there will be one  $\mathcal{M}'$  falsifying  $\Phi$ .

### 3.1 Model Contraction of Executability Laws

To contract an executability law  $\varphi \rightarrow \langle a \rangle \top$  from *one* model, one intuitively *removes arrows* leaving  $\varphi$ -worlds. In order to succeed in the operation, we have to guarantee that in the resulting model there will be at least one  $\varphi$ -world with no departing  $a$ -arrow.

**Definition 3.1** Let  $\mathcal{M} = \langle W, R \rangle$ .  $\mathcal{M}' = \langle W', R' \rangle \in \mathcal{M}_{\varphi \rightarrow \langle a \rangle \top}^{-}$  if and only if

- $W' = W$
- $R' \subseteq R$
- If  $(w, w') \in R \setminus R'$ , then  $\models_w^{\mathcal{M}} \varphi$
- There is  $w \in W'$  s.t.  $\not\models_w^{\mathcal{M}'} \varphi \rightarrow \langle a \rangle \top$

Observe that  $\mathcal{M}_{\varphi \rightarrow \langle a \rangle \top}^{-} \neq \emptyset$  if and only if  $\varphi$  is satisfiable in  $W$ . Moreover,  $\mathcal{M} \in \mathcal{M}_{\varphi \rightarrow \langle a \rangle \top}^{-}$  if and only if  $\not\models^{\mathcal{M}} \varphi \rightarrow \langle a \rangle \top$ .

To get minimal change, we want such an operation to be minimal w.r.t. the original model: one should remove a minimum set of arrows sufficient to get the desired result.

**Definition 3.2**  $\text{contract}(\mathcal{M}, \varphi \rightarrow \langle a \rangle \top) = \bigcup \min\{\mathcal{M}_{\varphi \rightarrow \langle a \rangle \top}^{-}, \preceq_{\mathcal{M}}\}$

And now we define the sets of possible models resulting from the contraction of an executability law in a set of models:

**Definition 3.3** Let  $\mathcal{M}$  be a set of models, and  $\varphi \rightarrow \langle a \rangle \top$  an executability law. Then

$$\mathcal{M}_{\varphi \rightarrow \langle a \rangle \top}^- = \{\mathcal{M}' : \mathcal{M}' = \mathcal{M} \cup \{\mathcal{M}'\}, \mathcal{M}' \in \text{contract}(\mathcal{M}, \varphi \rightarrow \langle a \rangle \top), \mathcal{M} \in \mathcal{M}\}$$

In our running example, consider  $\mathcal{M} = \{\mathcal{M}\}$ , where  $\mathcal{M}$  is the model in Figure 4. When the agent discovers that even with a token she does not manage to buy a coffee any more, she has to change her models in order to admit (new) models with states where *token* is the case but from which there is no *buy*-transition at all. Because having just one such world in each new model is enough, taking those resulting models whose accessibility relations are maximal guarantees minimal change. Hence we will have  $\mathcal{M}_{\text{token} \rightarrow \langle \text{buy} \rangle \top}^- = \{\mathcal{M} \cup \{\mathcal{M}'_1\}, \mathcal{M} \cup \{\mathcal{M}'_2\}, \mathcal{M} \cup \{\mathcal{M}'_3\}\}$ , where each  $\mathcal{M}'_i$  is depicted in Figure 5.

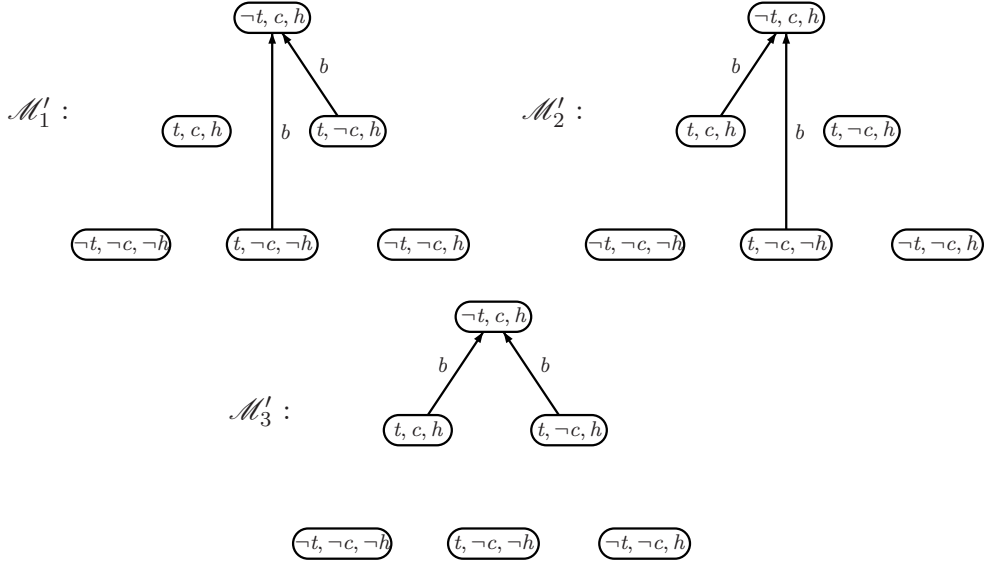


Figure 5: Models resulting from contracting  $\text{token} \rightarrow \langle \text{buy} \rangle \top$  in the model  $\mathcal{M}$  of Figure 4.

Clearly, if  $\varphi$  is not satisfied in  $\mathcal{M}$ , i.e.,  $\models^{\mathcal{M}} \neg \varphi$  for all  $\mathcal{M} \in \mathcal{M}$ , then the contraction of  $\varphi \rightarrow \langle a \rangle \top$  does not succeed. In this case,  $\neg \varphi$  should be contracted from the set of models (see further in this section).

### 3.2 Model Contraction of Effect Laws

When the agent discovers that there may be cases where after buying she gets no hot beverage, she must e.g. give up the belief  $token \rightarrow [buy]hot$  in her set of models. This means that  $token \wedge \langle buy \rangle \neg hot$  shall now be admitted in at least one world of some of her new models of beliefs. Hence, to contract an effect law  $\varphi \rightarrow [a]\psi$  from a given model, intuitively we have to *add arrows* leaving  $\varphi$ -worlds to worlds satisfying  $\neg\psi$ . The challenge in such an operation is how to guarantee minimal change.

In our example, when contracting  $token \rightarrow [buy]hot$  in the model of Figure 4 we add arrows from *token*-worlds to  $\neg hot$ -worlds. Because  $coffee \rightarrow hot$ , and then  $\neg hot \rightarrow \neg coffee$ , this should also give  $\langle buy \rangle \neg coffee$  in some *token*-world ( $\neg coffee$  is *relevant* to  $\neg hot$ , i.e., to have  $\neg hot$  we must have  $\neg coffee$ ). This means that if we allow for  $\langle buy \rangle \neg hot$  in some *token*-world, we also have to allow for  $\langle buy \rangle \neg coffee$  in that same world.

Hence, in our example one can add arrows from *token*-worlds to  $\neg hot \wedge \neg coffee \wedge token$ -worlds, as well as to  $\neg hot \wedge \neg coffee \wedge \neg token$  (Figure 6). For instance, one can add a *buy*-arrow from  $\{token, \neg coffee, \neg hot\}$  to one of these candidates (Figure 7).

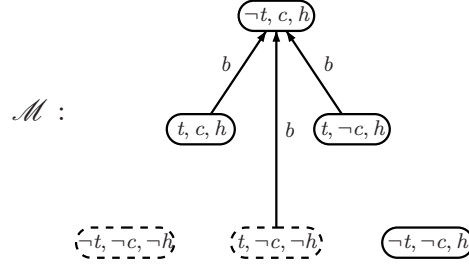


Figure 6: Candidate worlds to receive arrows from *token*-worlds.

Notice that adding the arrow to  $\{token, \neg coffee, \neg hot\}$  itself would make us lose the effect  $\neg token$ , true after every execution of *buy* in the original model ( $\models token \rightarrow [buy]\neg token$ ). How do we preserve this law while allowing for the new transition to a  $\neg hot$ -world? That is, how do we get rid of the effect *hot* without losing effects that are not relevant for that? We here develop an approach for this issue.

When adding a new arrow leaving a world  $w$  we intuitively want to preserve as many effects as we had before doing so. To achieve this, it is enough to preserve old effects only in  $w$  (because the remaining structure of



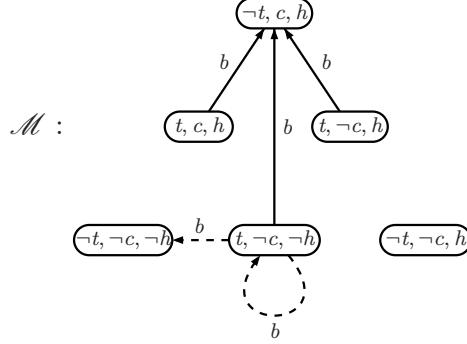


Figure 7: Two candidate new *buy*-arrows to falsify  $token \rightarrow [buy]hot$  in  $\mathcal{M}$ .

the model remains unchanged after adding the new arrow). Of course, we cannot preserve effects that are inconsistent with  $\neg\psi$  (those will all be lost). So, it suffices to preserve only the effects that are consistent with  $\neg\psi$ . To achieve that we must observe what is true in  $w$  and in the target world  $w'$ :

- What changes from  $w$  to  $w'$  ( $w' \setminus w$ ) must be what is obliged to do so: either because that is necessary to having  $\neg\psi$  in  $w'$  or because that is necessary to having another effect (independent of  $\neg\psi$ ) in  $w'$  that we want to preserve.
- What does not change from  $w$  to  $w'$  ( $w \cap w'$ ) should be what is allowed to do so: certain literals are never preserved (like *token* in our example), then when pointing the arrow to a world where it does not change w.r.t. the leaving world ( $\neg hot \wedge \neg coffee \wedge token$  in our example), we lose effects that held in  $w$  before adding the arrow.

This means that the only things allowed to change in the candidate target world must be those that are forced to change, either by some non-related law or because of having  $\neg\psi$  modulo a set of states  $W$ . In other words, we want the literals that change to be *at most* those that are sufficient to get  $\neg\psi$  modulo  $W$ , while preserving the maximum of effects. Every change outside that is not an intended one. Similarly, we want the literals that are preserved in the target world to be *at most* those that are usually preserved in a given set of models. Every preservation outside those may make us lose some law. This looks like prime implicants, and that is where prime subvaluations play their role: the worlds to which the new arrow will point are those whose difference w.r.t. the departing world are literals that are relevant and whose similarity w.r.t. it are literals that we know do not change.

**Definition 3.4 (Relevant Target Worlds)** Let  $\mathcal{M} = \langle W, R \rangle$  be a model,  $w, w' \in W$ ,  $\mathcal{M}$  a set of models such that  $\mathcal{M} \in \mathcal{M}$ , and  $\varphi \rightarrow [a]\psi$  an effect law. Then  $w'$  is a relevant target world of  $w$  w.r.t.  $\varphi \rightarrow [a]\psi$  for  $\mathcal{M}$  in  $\mathcal{M}$  if and only if

- $\models_w^{\mathcal{M}} \varphi, \not\models_{w'}^{\mathcal{M}} \psi$
- for all  $\ell \in w' \setminus w$ 
  - either there is  $v \in \text{base}(\neg\psi, W)$  s.t.  $v \subseteq w'$  and  $\ell \in v$
  - or there is  $\psi' \in \mathfrak{Fml}$  s.t. there is  $v' \in \text{base}(\psi', W)$  s.t.  $v' \subseteq w'$ ,  $\ell \in v'$ , and for every  $\mathcal{M}_i \in \mathcal{M}$ ,  $\models_w^{\mathcal{M}_i} [a]\psi'$
- for all  $\ell \in w \cap w'$ 
  - either there is  $v \in \text{base}(\neg\psi, W)$  s.t.  $v \subseteq w'$  and  $\ell \in v$
  - or there is  $\mathcal{M}_i \in \mathcal{M}$  such that  $\not\models_w^{\mathcal{M}_i} [a]\neg\ell$

By  $\text{RelTarget}(w, \varphi \rightarrow [a]\psi, \mathcal{M}, \mathcal{M})$  we denote the set of all relevant target worlds of  $w$  w.r.t.  $\varphi \rightarrow [a]\psi$  for  $\mathcal{M}$  in  $\mathcal{M}$ .

Note that we need the set of models  $\mathcal{M}$  (and here we can suppose it contains all models of the theory we want to change) because preserving effects depends on what other effects hold in the other models that interest us. We need to take them into account in the local operation of changing one model:<sup>8</sup>

**Definition 3.5** Let  $\mathcal{M} = \langle W, R \rangle$ , and  $\mathcal{M}$  be such that  $\mathcal{M} \in \mathcal{M}$ . Then  $\mathcal{M}' = \langle W', R' \rangle \in \mathcal{M}_{\varphi \rightarrow [a]\psi}^-$  if and only if

- $W' = W$
- $R \subseteq R'$
- If  $(w, w') \in R' \setminus R$ , then  $w' \in \text{RelTarget}(w, \varphi \rightarrow [a]\psi, \mathcal{M}, \mathcal{M})$
- There is  $w \in W'$  s.t.  $\not\models_w^{\mathcal{M}'} \varphi \rightarrow [a]\psi$

---

<sup>8</sup>The reason we do not need  $\mathcal{M}$  in the definition of the local (one model) contraction of executability laws  $\mathcal{M}_{\varphi \rightarrow (a)\top}^-$  is that when removing arrows there is no way of losing effects, as every effect law that held in the world from which an arrow has been removed remains true in the same world in the resulting model.

Observe that  $\mathcal{M}_{\varphi \rightarrow [a]\psi}^- \neq \emptyset$  if and only if  $\varphi$  and  $\neg\psi$  are both satisfiable in  $W$ . Moreover,  $\mathcal{M} \in \mathcal{M}_{\varphi \rightarrow [a]\psi}^-$  if and only if  $\not\models^{\mathcal{M}} \varphi \rightarrow [a]\psi$ .

Because having just one world where the law is no longer true in each model is enough, taking those resulting models whose accessibility relations are minimal w.r.t. the original one guarantees minimal change.

**Definition 3.6**  $contract(\mathcal{M}, \varphi \rightarrow [a]\psi) = \bigcup \min\{\mathcal{M}_{\varphi \rightarrow [a]\psi}^-, \preceq \mathcal{M}\}$

Now we can define the possible sets of models resulting from contracting an effect law from a set of models:

**Definition 3.7** Let  $\mathcal{M}$  be a set of models, and  $\varphi \rightarrow [a]\psi$  an effect law. Then

$$\mathcal{M}_{\varphi \rightarrow [a]\psi}^- = \{\mathcal{M}' : \mathcal{M}' = \mathcal{M} \cup \{\mathcal{M}'\}, \mathcal{M}' \in contract(\mathcal{M}, \varphi \rightarrow [a]\psi), \mathcal{M} \in \mathcal{M}\}$$

Taking again  $\mathcal{M} = \{\mathcal{M}\}$ , where  $\mathcal{M}$  is the model in Figure 4, after contracting  $token \rightarrow [buy]hot$  from  $\mathcal{M}$  we get  $\mathcal{M}_{token \rightarrow [buy]hot}^- = \{\mathcal{M} \cup \{\mathcal{M}'_1\}, \mathcal{M} \cup \{\mathcal{M}'_2\}, \mathcal{M} \cup \{\mathcal{M}'_3\}\}$ , where all  $\mathcal{M}'_i$ 's are as depicted in Figure 8.

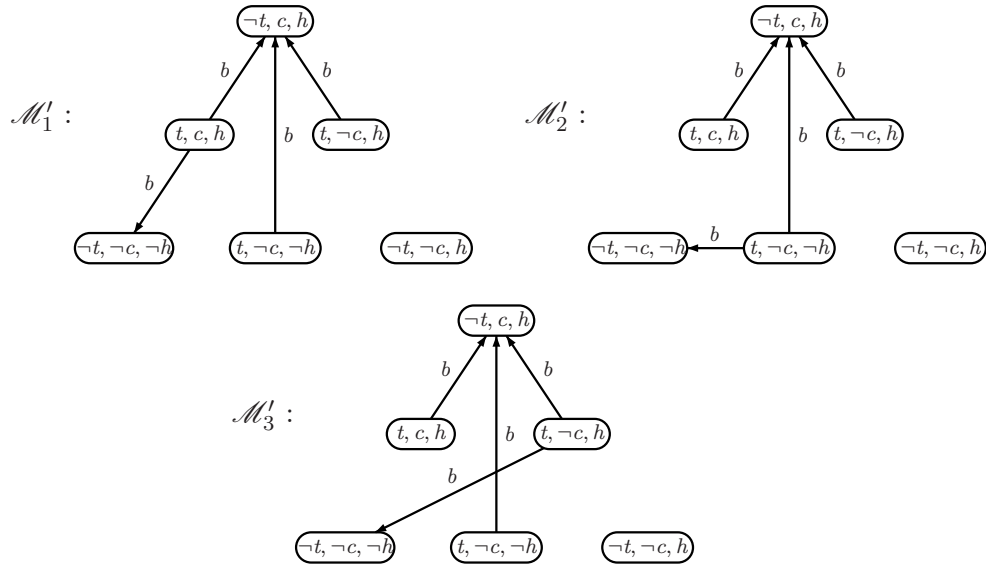


Figure 8: Models resulting from contracting  $token \rightarrow [buy]hot$  in the model  $\mathcal{M}$  of Figure 4.

In both cases where  $\varphi$  is not satisfiable in  $\mathcal{M}$  or  $\psi$  is valid in  $\mathcal{M}$ , of course our operator does not succeed in falsifying  $\varphi \rightarrow [a]\psi$  (cf. end of Section 3.1). Intuitively, prior to doing that we have to change our set of possible states. This is what we address in the next section.

### 3.3 Model Contraction of Static Laws

When contracting a static law from a model, we want to admit the existence of at least one (new) possible state falsifying it. This means that intuitively we should *add new worlds* to the original model. This is quite easy. A very delicate issue however is what to do with the accessibility relation: should new arrows leave/arrive at the new world? If no arrow leaves the new added world, we may lose some executability law. If some arrow leaves it, then we may lose some effect law, the same holding if we add an arrow pointing to the new world. On the other hand, if no arrow arrives at the new world, what about the intuition? Is it intuitive to have an unreachable state?

All this discussion shows how drastic a change in the static laws may be: it is a change in the underlying structure (possible states) of the world! Changing it may have as consequence the loss of an effect law or an executability law. What we can do is choose which laws we accept to lose and postpone their change (by the other operators). Following the tradition in the RAA community which states that executability laws are, in general, more difficult to formalize than effect laws, and hence are more likely to be incorrect, here we prefer not to change the accessibility relation, which means preserving effect laws and postponing correction of executability laws, if needed. (cf. Sections 4.3 and 10 below).

**Definition 3.8** Let  $\mathcal{M} = \langle W, R \rangle$ .  $\mathcal{M}' = \langle W', R' \rangle \in \mathcal{M}_\varphi^-$  if and only if

- $W \subseteq W'$
- $R = R'$
- There is  $w \in W'$  s.t.  $\not\models_w^{\mathcal{M}'} \varphi$

Notice that we have  $\mathcal{M}_\varphi^- = \emptyset$  if and only if  $\varphi$  is a tautology. Moreover,  $\mathcal{M} \in \mathcal{M}_\varphi^-$  if and only if  $\not\models^{\mathcal{M}} \varphi$ .

The minimal modifications of one model are defined as usual:

**Definition 3.9**  $\text{contract}(\mathcal{M}, \varphi) = \bigcup \min\{\mathcal{M}_\varphi^-, \preceq_{\mathcal{M}}\}$

And now we define the sets of models resulting from contracting a static law from a given set of models:

**Definition 3.10** *Let  $\mathcal{M}$  be a set of models, and  $\varphi$  a static law. Then*

$$\mathcal{M}_{\varphi}^{-} = \{\mathcal{M}' : \mathcal{M}' = \mathcal{M} \cup \{\mathcal{M}'\}, \mathcal{M}' \in \text{contract}(\mathcal{M}, \varphi), \mathcal{M} \in \mathcal{M}\}$$

In our scenario example, if  $\mathcal{M} = \{\mathcal{M}\}$ , where  $\mathcal{M}$  is the model in Figure 4, then contracting  $\text{coffee} \rightarrow \text{hot}$  from  $\mathcal{M}$  would give us  $\mathcal{M}_{\text{coffee} \rightarrow \text{hot}}^{-} = \{\mathcal{M} \cup \{\mathcal{M}'_1\}, \mathcal{M} \cup \{\mathcal{M}'_2\}\}$ , where each  $\mathcal{M}'_i$  is as depicted in Figure 9.

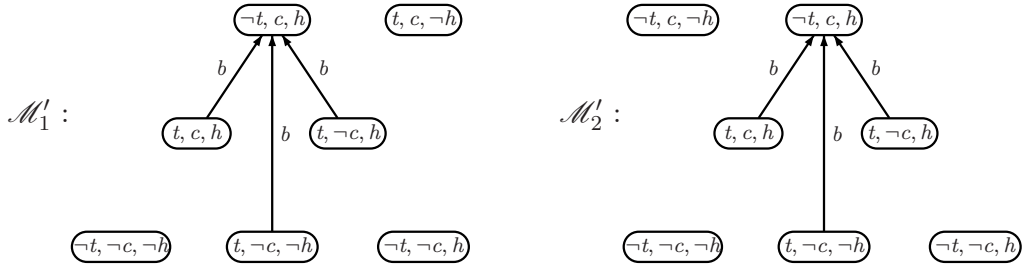


Figure 9: Models resulting from contracting  $\text{coffee} \rightarrow \text{hot}$  in the model  $\mathcal{M}$  of Figure 4.

Notice that by not modifying the accessibility relation all the effect laws are preserved with minimal change. Moreover, our approach is in line with intuition: when learning that a new state is now possible, we do not necessarily know all the behavior of the actions in the new added state. We may expect some action laws to hold in the new state (see Section 10 for an alternative solution), but, with the information we dispose, not touching the accessibility relation is the safest way of contracting static laws.

## 4 Syntactic Operators for Contraction of Laws

Now that we have defined the semantics of our theory change, we turn our attention to the definition of syntactic operators for changing sets of formulas.

As Nebel [40] says, “[...] finite bases usually represent [...] laws, and when we are forced to change the theory we would like to stay as close as possible to the original [...] base.” Hence, besides the definition of syntactical operators, we should also guarantee that they perform minimal change.

By  $\mathcal{T}_{\Phi}^-$  we denote in the sequel the result of contracting a law  $\Phi$  from the set of laws  $\mathcal{T}$ .

#### 4.1 Contracting Executability Laws

For the case of contracting  $\varphi \rightarrow \langle a \rangle \top$  from an action theory, first we have to ensure that the action  $a$  is still executable in all those contexts where  $\neg\varphi$  is the case. Second, in order to get minimality, we must make  $a$  executable in *some* contexts where  $\varphi$  is true, viz. all  $\varphi$ -worlds but one. This means that we can have several action theories as outcome. Algorithm 1 gives a syntactical operator to achieve this.

---

**Algorithm 1** Erasure of an executability law

---

**input:**  $\mathcal{T}, \varphi \rightarrow \langle a \rangle \top$

**output:**  $\mathcal{T}_{\varphi \rightarrow \langle a \rangle \top}^-$

1:  $\mathcal{T}_{\varphi \rightarrow \langle a \rangle \top}^- := \emptyset$

2: **if**  $\mathcal{T} \models_{\text{PDL}} \varphi \rightarrow \langle a \rangle \top$  **then**

3:   **for all**  $\pi \in IP(\mathcal{S} \wedge \varphi)$  **do**

4:     **for all**  $A \subseteq \overline{atm(\pi)}$  **do**

5:        $\varphi_A := \bigwedge_{\substack{p_i \in atm(\pi) \\ p_i \in A}} p_i \wedge \bigwedge_{\substack{p_i \in atm(\pi) \\ p_i \notin A}} \neg p_i$

6:       **if**  $\mathcal{S} \not\models_{\text{CPL}} (\pi \wedge \varphi_A) \rightarrow \perp$  **then**

7:          $\mathcal{T}' := \left( (\mathcal{T} \setminus \mathcal{X}_a) \cup \{(\varphi_i \wedge \neg(\pi \wedge \varphi_A)) \rightarrow \langle a \rangle \top : \varphi_i \rightarrow \langle a \rangle \top \in \mathcal{X}_a\} \right)$

8:        $\mathcal{T}_{\varphi \rightarrow \langle a \rangle \top}^- := \mathcal{T}_{\varphi \rightarrow \langle a \rangle \top}^- \cup \{\mathcal{T}'\}$

9: **else**

10:    $\mathcal{T}_{\varphi \rightarrow \langle a \rangle \top}^- := \{\mathcal{T}\}$

**return**  $\mathcal{T}_{\varphi \rightarrow \langle a \rangle \top}^-$

---

Observe that from the finiteness of  $\mathcal{T}$  and that of  $\overline{atm(\pi)}$ , for any  $\pi \in IP(\mathcal{S} \wedge \varphi)$ , and the decidability of PDL [20] and of classical propositional logic, it follows that Algorithm 1 terminates.

In our running example, contracting the executability law  $token \rightarrow \langle buy \rangle \top$  from the action theory  $\mathcal{T}$  would give us  $\mathcal{T}_{token \rightarrow \langle buy \rangle \top}^- = \{\mathcal{T}'_1, \mathcal{T}'_2, \mathcal{T}'_3\}$ ,

where:

$$\begin{aligned}
\mathcal{T}'_1 &= \left\{ \begin{array}{l} \text{coffee} \rightarrow \text{hot}, \neg \text{coffee} \rightarrow [\text{buy}] \text{coffee}, \\ \text{token} \rightarrow [\text{buy}] \neg \text{token}, \neg \text{token} \rightarrow [\text{buy}] \perp, \\ \text{coffee} \rightarrow [\text{buy}] \text{coffee}, \text{hot} \rightarrow [\text{buy}] \text{hot}, \\ (\text{token} \wedge \neg \text{coffee} \wedge \text{hot}) \rightarrow \langle \text{buy} \rangle \top, \\ (\text{token} \wedge \neg \text{coffee} \wedge \neg \text{hot}) \rightarrow \langle \text{buy} \rangle \top \end{array} \right\} \\
\mathcal{T}'_2 &= \left\{ \begin{array}{l} \text{coffee} \rightarrow \text{hot}, \neg \text{coffee} \rightarrow [\text{buy}] \text{coffee}, \\ \text{token} \rightarrow [\text{buy}] \neg \text{token}, \neg \text{token} \rightarrow [\text{buy}] \perp, \\ \text{coffee} \rightarrow [\text{buy}] \text{coffee}, \text{hot} \rightarrow [\text{buy}] \text{hot}, \\ (\text{token} \wedge \text{coffee} \wedge \text{hot}) \rightarrow \langle \text{buy} \rangle \top, \\ (\text{token} \wedge \neg \text{coffee} \wedge \neg \text{hot}) \rightarrow \langle \text{buy} \rangle \top \end{array} \right\} \\
\mathcal{T}'_3 &= \left\{ \begin{array}{l} \text{coffee} \rightarrow \text{hot}, \neg \text{coffee} \rightarrow [\text{buy}] \text{coffee}, \\ \text{token} \rightarrow [\text{buy}] \neg \text{token}, \neg \text{token} \rightarrow [\text{buy}] \perp, \\ \text{coffee} \rightarrow [\text{buy}] \text{coffee}, \text{hot} \rightarrow [\text{buy}] \text{hot}, \\ (\text{token} \wedge \text{coffee} \wedge \text{hot}) \rightarrow \langle \text{buy} \rangle \top, \\ (\text{token} \wedge \neg \text{coffee} \wedge \text{hot}) \rightarrow \langle \text{buy} \rangle \top \end{array} \right\}
\end{aligned}$$

Now the knowledge engineer has only to choose which theory is more in line with her intuitions and implement the changes (cf. Figure 5).

## 4.2 Contracting Effect Laws

When contracting  $\varphi \rightarrow [a]\psi$  from a theory  $\mathcal{T}$ , intuitively we should contract some effect laws that preclude  $\neg\psi$  in target worlds. In order to cope with minimality, we must change only those laws that are relevant to  $\varphi \rightarrow [a]\psi$ .

Let  $\mathcal{E}_a^{\varphi, \psi}$  denote a minimum subset of  $\mathcal{E}_a$  such that  $\mathcal{S}, \mathcal{E}_a^{\varphi, \psi} \models_{\text{PDL}} \varphi \rightarrow [a]\psi$ . In the case the theory is modular [25] (see further), such a set always exists. Moreover, note that there can be more than one such a set, in which case we denote them  $(\mathcal{E}_a^{\varphi, \psi})_1, \dots, (\mathcal{E}_a^{\varphi, \psi})_n$ . Let

$$\mathcal{E}_a^- = \bigcup_{1 \leq i \leq n} (\mathcal{E}_a^{\varphi, \psi})_i$$

The laws in  $\mathcal{E}_a^-$  will serve as guidelines to get rid of  $\varphi \rightarrow [a]\psi$  in the theory.

The first thing we must do is to ensure that action  $a$  still has effect  $\psi$  in all those contexts in which  $\varphi$  does not hold. This means we shall weaken the laws in  $\mathcal{E}_a^{\varphi, \psi}$  specializing them to  $\neg\varphi$ . Now, we need to preserve all old effects in all  $\varphi$ -worlds but one. To achieve that we specialize the above

laws to each possible valuation (maximal conjunction of literals) satisfying  $\varphi$  but one. Then, in the left  $\varphi$ -valuation, we must ensure that action  $a$  has either its old effects or  $\neg\psi$  as outcome. We achieve that by weakening the *consequent* of the laws in  $\mathcal{E}_a^-$ . Finally, in order to get minimal change, we must ensure that all literals in this  $\varphi$ -valuation that are not forced to change in  $\neg\psi$ -worlds should be preserved. We do this by stating an effect law of the form  $(\varphi_k \wedge \ell) \rightarrow [a](\psi \vee \ell)$ , where  $\varphi_k$  is the above  $\varphi$ -valuation. The reason this is needed is clear: there can be several  $\neg\psi$ -valuations, and as far as we want at most one to be reachable from the  $\varphi_k$ -world, we should force it to be the one whose difference to this  $\varphi_k$ -valuation is minimal.

Again, the result will be a set of action theories. Algorithm 2 below gives the operator.

The reader is invited to check that Algorithm 2 always terminates (cf. Section 4.1).

For an example of execution of the algorithm, suppose we want to contract the effect law  $token \rightarrow [buy]hot$  from our theory  $\mathcal{T}$ . We first determine the minimum sets of effect laws that together with  $\mathcal{S}$  entail  $token \rightarrow [buy]hot$ . They are

$$\begin{aligned} (\mathcal{E}_{buy}^{token,hot})_1 &= \left\{ \begin{array}{l} coffee \rightarrow [buy]coffee, \\ \neg coffee \rightarrow [buy]coffee \end{array} \right\} \\ (\mathcal{E}_{buy}^{token,hot})_2 &= \left\{ \begin{array}{l} hot \rightarrow [buy]hot, \\ \neg coffee \rightarrow [buy]coffee \end{array} \right\} \end{aligned}$$

Now for each context where  $token$  is the case, we weaken the effect laws in  $\mathcal{E}_{buy}^- = (\mathcal{E}_{buy}^{token,hot})_1 \cup (\mathcal{E}_{buy}^{token,hot})_2$ . Given  $\mathcal{S} = \{coffee \rightarrow hot\}$ , such contexts are  $token \wedge coffee \wedge hot$ ,  $token \wedge \neg coffee \wedge \neg hot$  and  $token \wedge \neg coffee \wedge hot$ .

For  $token \wedge coffee \wedge hot$ : we replace in  $\mathcal{T}$  the laws from  $\mathcal{E}_{buy}^-$  by

$$\left\{ \begin{array}{l} (coffee \wedge \neg(token \wedge coffee \wedge hot)) \rightarrow [buy]coffee, \\ (hot \wedge \neg(token \wedge coffee \wedge hot)) \rightarrow [buy]hot, \\ (\neg coffee \wedge \neg(token \wedge coffee \wedge hot)) \rightarrow [buy]coffee \end{array} \right\}$$

so that we preserve their effects in all possible contexts but  $token \wedge coffee \wedge hot$ . Now, in order to preserve some effects in  $token \wedge coffee \wedge hot$ -contexts while allowing for reachable  $\neg hot$ -worlds, we add the laws:

$$\left\{ \begin{array}{l} (token \wedge coffee \wedge hot) \rightarrow [buy](coffee \vee \neg hot), \\ (token \wedge coffee \wedge hot) \rightarrow [buy](hot \vee \neg coffee) \end{array} \right\}$$



---

**Algorithm 2** Contraction of an effect law
 

---

**input:**  $\mathcal{T}, \varphi \rightarrow [a]\psi$

**output:**  $\mathcal{T}_{\varphi \rightarrow [a]\psi}^-$

1:  $\mathcal{T}_{\varphi \rightarrow [a]\psi}^- := \emptyset$

2: **if**  $\mathcal{T} \models_{\text{PDL}} \varphi \rightarrow [a]\psi$  **then**

3:   **for all**  $\pi \in IP(\mathcal{S} \wedge \varphi)$  **do**

4:     **for all**  $A \subseteq \overline{atm(\pi)}$  **do**

5:        $\varphi_A := \bigwedge_{\substack{p_i \in \overline{atm(\pi)} \\ p_i \in A}} p_i \wedge \bigwedge_{\substack{p_i \in \overline{atm(\pi)} \\ p_i \notin A}} \neg p_i$

6:       **if**  $\mathcal{S} \not\models_{\text{CPL}} (\pi \wedge \varphi_A) \rightarrow \perp$  **then**

7:          **for all**  $\pi' \in IP(\mathcal{S} \wedge \neg\psi)$  **do**

8:           
$$\mathcal{T}' := \left( \begin{array}{l} (T \setminus \mathcal{E}_a^-) \cup \\ \{(\varphi_i \wedge \neg(\pi \wedge \varphi_A)) \rightarrow [a]\psi_i : \varphi_i \rightarrow [a]\psi_i \in \mathcal{E}_a^-\} \cup \\ \{(\varphi_i \wedge \pi \wedge \varphi_A) \rightarrow [a](\psi_i \vee \pi') : \varphi_i \rightarrow [a]\psi_i \in \mathcal{E}_a^-\} \end{array} \right)$$

9:       **for all**  $L \subseteq \mathfrak{L}$  **do**

10:          **if**  $\mathcal{S} \models_{\text{CPL}} (\pi \wedge \varphi_A) \rightarrow \bigwedge_{\ell \in L} \ell$  **and**  $\mathcal{S} \not\models_{\text{CPL}} (\pi' \wedge \bigwedge_{\ell \in L} \ell) \rightarrow \perp$  **then**

11:           **for all**  $\ell \in L$  **do**

12:             **if**  $\mathcal{T} \not\models_{\text{PDL}} (\pi \wedge \varphi_A \wedge \ell) \rightarrow [a]\neg\ell$  **or**  $\ell \in \pi'$  **then**

13:                $\mathcal{T}' := \mathcal{T}' \cup \{(\pi \wedge \varphi_A \wedge \ell) \rightarrow [a](\psi \vee \ell)\}$

14:            $\mathcal{T}_{\varphi \rightarrow [a]\psi}^- := \mathcal{T}_{\varphi \rightarrow [a]\psi}^- \cup \{\mathcal{T}'\}$

15: **else**

16:    $\mathcal{T}_{\varphi \rightarrow [a]\psi}^- := \{\mathcal{T}\}$

**return**  $\mathcal{T}_{\varphi \rightarrow [a]\psi}^-$

---

Now, we search all possible combinations of laws from  $\mathcal{E}_{buy}$  that apply on  $token \wedge coffee \wedge hot$  contexts and find  $token \rightarrow [buy]\neg token$ . Because  $\neg token$  must be true after every execution of  $buy$ , we do not state the law  $(token \wedge coffee \wedge hot) \rightarrow [buy](hot \vee token)$ , and end up with the theory:

$$\mathcal{T}'_1 = \left\{ \begin{array}{l} coffee \rightarrow hot, token \rightarrow \langle buy \rangle \top, \\ token \rightarrow [buy]\neg token, \neg token \rightarrow [buy]\perp, \\ (coffee \wedge \neg(token \wedge coffee \wedge hot)) \rightarrow [buy]coffee, \\ (hot \wedge \neg(token \wedge coffee \wedge hot)) \rightarrow [buy]hot, \\ (\neg coffee \wedge \neg(token \wedge coffee \wedge hot)) \rightarrow [buy]coffee, \\ (token \wedge coffee \wedge hot) \rightarrow [buy](coffee \vee \neg hot), \\ (token \wedge coffee \wedge hot) \rightarrow [buy](hot \vee \neg coffee) \end{array} \right\}$$

On the other hand, if in our language we also had an atom  $p$  with the same theory  $\mathcal{T}$ , then we should add a law  $(token \wedge coffee \wedge hot \wedge p) \rightarrow [buy](hot \vee p)$  to meet minimal change by preserving effects that are not relevant to  $\neg\psi$ .

The execution for contexts  $token \wedge \neg coffee \wedge \neg hot$  and  $token \wedge \neg coffee \wedge hot$  are analogous and the algorithm ends with  $\mathcal{T}^-_{token \rightarrow [buy]hot} = \{\mathcal{T}'_1, \mathcal{T}'_2, \mathcal{T}'_3\}$ , where:

$$\mathcal{T}'_2 = \left\{ \begin{array}{l} coffee \rightarrow hot, token \rightarrow \langle buy \rangle \top, \\ token \rightarrow [buy]\neg token, \neg token \rightarrow [buy]\perp, \\ (coffee \wedge \neg(token \wedge \neg coffee \wedge \neg hot)) \rightarrow [buy]coffee, \\ (hot \wedge \neg(token \wedge \neg coffee \wedge \neg hot)) \rightarrow [buy]hot, \\ (\neg coffee \wedge \neg(token \wedge \neg coffee \wedge \neg hot)) \rightarrow [buy]coffee, \\ (token \wedge \neg coffee \wedge \neg hot) \rightarrow [buy](coffee \vee \neg hot) \end{array} \right\}$$

$$\mathcal{T}'_3 = \left\{ \begin{array}{l} coffee \rightarrow hot, token \rightarrow \langle buy \rangle \top, \\ token \rightarrow [buy]\neg token, \neg token \rightarrow [buy]\perp, \\ (coffee \wedge \neg(token \wedge \neg coffee \wedge hot)) \rightarrow [buy]coffee, \\ (hot \wedge \neg(token \wedge \neg coffee \wedge hot)) \rightarrow [buy]hot, \\ (\neg coffee \wedge \neg(token \wedge \neg coffee \wedge hot)) \rightarrow [buy]coffee, \\ (token \wedge \neg coffee \wedge hot) \rightarrow [buy](hot \vee \neg coffee), \\ (token \wedge \neg coffee \wedge hot) \rightarrow [buy](coffee \vee \neg hot) \end{array} \right\}$$

Looking at Figure 8, we can see the correspondence between these theories and their respective models.

### 4.3 Contracting Static Laws

Finally, in order to contract a static law from a theory, we can use any contraction/erasure operator  $\ominus$  for classical logic. Because contracting static laws means *admitting* new possible states (cf. the semantics), just modifying the set  $\mathcal{S}$  of static laws may not be enough for the dynamic logic case. Since we in general do not necessarily know the behavior of the actions in a new discovered state of the world, a careful approach is to change the theory so that all action laws remain the same in the contexts where the contracted law is the case. In our example, if when contracting the law  $coffee \rightarrow hot$  we are not sure whether *buy* is still executable or not, then we should weaken our executability laws specializing them to the context  $coffee \rightarrow hot$ , and make *buy* a priori inexecutable in all  $\neg(coffee \rightarrow hot)$  contexts. The operator given in Algorithm 3 formalizes this.

---

**Algorithm 3** Contraction of a static law

---

**input:**  $\mathcal{T}, \varphi$

**output:**  $\mathcal{T}_\varphi^-$

1:  $\mathcal{T}_\varphi^- := \emptyset$

2: **if**  $\mathcal{S} \models_{\text{CPL}} \varphi$  **then**

3:   **for all**  $\mathcal{S}^- \in \mathcal{S} \ominus \varphi$  **do**

4:      $\mathcal{T}' := \left( \begin{array}{c} ((\mathcal{T} \setminus \mathcal{S}) \cup \mathcal{S}^-) \setminus \mathcal{X}_a \cup \\ \{(\varphi_i \wedge \varphi) \rightarrow \langle a \rangle \top : \varphi_i \rightarrow \langle a \rangle \top \in \mathcal{X}_a\} \cup \\ \{\neg \varphi \rightarrow [a] \perp\} \end{array} \right)$

5:      $\mathcal{T}_\varphi^- := \mathcal{T}_\varphi^- \cup \{\mathcal{T}'\}$

6: **else**

7:    $\mathcal{T}_\varphi^- := \{\mathcal{T}\}$

**return**  $\mathcal{T}_\varphi^-$

---

In our running example, contracting the law  $coffee \rightarrow hot$  from  $\mathcal{T}$  produces  $\mathcal{T}_{coffee \rightarrow hot}^- = \{\mathcal{T}'_1, \mathcal{T}'_2\}$ , where

$$\mathcal{T}'_1 = \left\{ \begin{array}{l} \neg(\neg token \wedge coffee \wedge \neg hot), \\ (token \wedge coffee \rightarrow hot) \rightarrow \langle buy \rangle \top, \\ \neg coffee \rightarrow [buy] coffee, token \rightarrow [buy] \neg token, \\ \neg token \rightarrow [buy] \perp, coffee \rightarrow [buy] coffee, \\ hot \rightarrow [buy] hot, (coffee \wedge \neg hot) \rightarrow [buy] \perp \end{array} \right\}$$

$$\mathcal{T}'_2 = \left\{ \begin{array}{l} \neg(token \wedge coffee \wedge \neg hot), \\ (token \wedge coffee \rightarrow hot) \rightarrow \langle buy \rangle \top, \\ \neg coffee \rightarrow [buy] coffee, token \rightarrow [buy] \neg token, \\ \neg token \rightarrow [buy] \perp, coffee \rightarrow [buy] coffee, \\ hot \rightarrow [buy] hot, (coffee \wedge \neg hot) \rightarrow [buy] \perp \end{array} \right\}$$

Observe that the effect laws are not affected by the change: as far as we do not pronounce ourselves about the executability of some action in the new added world, all the effect laws remain true in it.

If the knowledge engineer is not happy with  $(coffee \wedge \neg hot) \rightarrow [buy] \perp$ , she can contract this formula from the theory using Algorithm 2. Ideally, besides stating that *buy* is executable in the context  $coffee \wedge \neg hot$ , we should want to specify its outcome in this context as well. For example, we could want  $(coffee \wedge \neg hot) \rightarrow \langle buy \rangle hot$  to be true in the result. This would require theory *revision*. See Section 7 for the semantics of such an operation.

## 5 Correctness of the Operators

We here address the correctness of our algorithms w.r.t. our semantics for contraction.

### 5.1 Two Counter-Examples

Let the theory  $\mathcal{T} = \{p_1 \rightarrow \langle a \rangle \top, (\neg p_1 \vee p_2) \rightarrow [a] \perp, [a] \neg p_2\}$  and consider its model  $\mathcal{M}$  depicted in Figure 10. (Notice that  $\mathcal{T} \models_{\text{PDL}} \neg(p_1 \wedge p_2)$ .) When contracting  $p_1 \rightarrow [a] \neg p_2$  in  $\mathcal{M}$ , we get  $\mathcal{M}'$  in Figure 10.

Now contracting  $p_1 \rightarrow [a] \neg p_2$  from  $\mathcal{T}$  using Algorithm 2 gives  $\mathcal{T}^-_{p_1 \rightarrow [a] \neg p_2} = \{\mathcal{T}'\}$ , where

$$\mathcal{T}' = \left\{ \begin{array}{l} p_1 \rightarrow \langle a \rangle \top, (\neg p_1 \vee p_2) \rightarrow [a] \perp, \\ (p_1 \wedge \neg p_2) \rightarrow [a] (\neg p_2 \vee p_2), \\ (p_1 \wedge \neg p_2) \rightarrow [a] (\neg p_2 \vee p_1) \end{array} \right\}$$

Notice that the formula  $(p_1 \wedge \neg p_2) \rightarrow [a] (\neg p_2 \vee p_1)$  is put in  $\mathcal{T}'$  by Algorithm 2 because there is  $\{p_1\} \subseteq \mathfrak{L}$  it such that  $\mathcal{S} \not\models_{\text{CPL}} (p_1 \wedge p_2) \rightarrow \perp$  and  $\mathcal{T} \not\models_{\text{PDL}}$

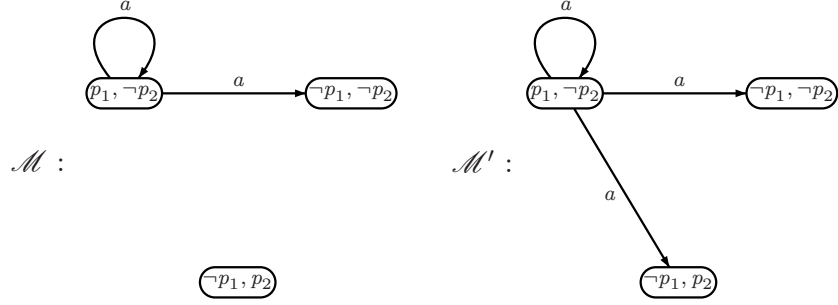


Figure 10: A model  $\mathcal{M}$  of  $\mathcal{T}$  and the result  $\mathcal{M}'$  of contracting  $p_1 \rightarrow [a]\neg p_2$  in it.

$(p_1 \wedge \neg p_2) \rightarrow [a]\neg p_1$ . Clearly  $\not\models^{\mathcal{M}'} \mathcal{T}'$  and no theory in  $\mathcal{T}_{p_1 \rightarrow [a]\neg p_2}^-$  has  $\mathcal{M}'$  as model. This means that the contraction operators are not correct.

This issue arises because Algorithm 2 tries to allow an arrow from the  $p_1 \wedge \neg p_2$ -world to a  $p_2$ -world that is closest to it, viz.  $\{p_1, p_2\}$ , but has no way of knowing that such a world does not exist. A remedy for that is replacing the test  $\mathcal{T} \not\models_{\text{DL}} (\pi' \wedge \bigwedge_{\ell \in L} \ell) \rightarrow \perp$  for  $\mathcal{S} \not\models_{\text{PL}} (\pi' \wedge \bigwedge_{\ell \in L} \ell) \rightarrow \perp$ , but that would increase even more the complexity of the algorithm. A better option would be to have  $\mathcal{S}$  ‘complete enough’ to allow the algorithm to determine the worlds to which a new transition could exist.

The other way round, it does not hold in general that the models of each  $\mathcal{T}' \in \mathcal{T}_{\Phi}^-$  result from the semantic contraction of models of  $\mathcal{T}$  by  $\Phi$ . To see this suppose that there is only one atom  $p$  and one action  $a$ , and consider the action theory  $\mathcal{T} = \{p \rightarrow [a]\perp, \langle a \rangle \top\}$ . The only model of  $\mathcal{T}$  is  $\mathcal{M} = \langle \{\{\neg p\}\}, \{(\{\neg p\}, \{\neg p\})\} \rangle$  in Figure 11.



Figure 11: Incompleteness of contraction: a model  $\mathcal{M}$  of  $\mathcal{T}$  and a model  $\mathcal{M}'$  of the theory resulting from contracting  $p \rightarrow \langle a \rangle \top$  from  $\mathcal{T}$ .

By definition,  $\text{contract}(\mathcal{M}, p \rightarrow \langle a \rangle \top) = \{\mathcal{M}\}$ . On the other hand,  $\mathcal{T}_{p \rightarrow \langle a \rangle \top}^-$  is the singleton  $\{\mathcal{T}'\}$  such that  $\mathcal{T}' = \{p \rightarrow [a]\perp, \neg p \rightarrow \langle a \rangle \top\}$ . Then

$\mathcal{M}' = \langle \{\{\neg p\}, \{p\}\}, (\{\neg p\}, \{\neg p\}) \rangle$  in Figure 11 is a model of the contracted theory. Clearly,  $\mathcal{M}'$  does not result from the semantic contraction of  $p \rightarrow \langle a \rangle \top$  from  $\mathcal{M}$ : while  $\neg p$  is valid in the contraction of the models of  $\mathcal{T}$ , it is not valid in the models of  $\mathcal{T}'$ . This means that the operators are not complete.

This problem occurs because, in our example, the worlds that are forbidden by  $\mathcal{T}$ , e.g.  $\{p\}$ , are not preserved as such in  $\mathcal{T}'$ . When contracting an executability or an effect law, we are not supposed to change the possible worlds of a theory (cf. Section 3).

Fortunately correctness of the algorithms w.r.t. our semantics can be guaranteed for those theories whose  $\mathcal{S}$  is maximal, i.e., the set of static laws in  $\mathcal{S}$  alone determine what worlds are authorized in the models of the theory. This is the principle of *modularity* [25] and we briefly review it in the next section.

## 5.2 Modular Theories

**Definition 5.1 (Modularity [25])** *An action theory  $\mathcal{T}$  is modular if and only if for every  $\varphi \in \mathfrak{Fml}$ , if  $\mathcal{T} \models_{\text{PDL}} \varphi$ , then  $\mathcal{S} \models_{\text{CPL}} \varphi$ .*

For an example of a non-modular theory, suppose that the action theory  $\mathcal{T}$  of our coffee machine scenario were stated as

$$\mathcal{T} = \left\{ \begin{array}{l} \text{coffee} \rightarrow \text{hot}, \underline{\langle \text{buy} \rangle \top}, \\ \neg \text{coffee} \rightarrow [\text{buy}] \text{coffee}, \\ \text{token} \rightarrow [\text{buy}] \neg \text{token}, \neg \text{token} \rightarrow [\text{buy}] \perp, \\ \text{coffee} \rightarrow [\text{buy}] \text{coffee}, \text{hot} \rightarrow [\text{buy}] \text{hot} \end{array} \right\}$$

The modified law is underlined: we have (in this case wrongly) stated that the agent can always buy at the machine. Then  $\mathcal{T} \models_{\text{PDL}} \text{token}$  and  $\mathcal{S} \not\models_{\text{CPL}} \text{token}$ .

As the underlying multimodal logic is independently axiomatized (see Section 2.1), we can use the algorithms given by Herzig and Varzinczak [25] to check whether an action theory satisfies the principle of modularity. Whenever this is not the case, the algorithms return the Boolean formulas entailed by the theory that are not consequences of  $\mathcal{S}$  alone. For the theory  $\mathcal{T}$  above, they would return  $\{\text{token}\}$ : as we stated  $\langle \text{buy} \rangle \top$ , from this and  $\neg \text{token} \rightarrow [\text{buy}] \perp$  we get  $\mathcal{T} \models_{\text{PDL}} \text{token}$ . Because  $\mathcal{S} \not\models_{\text{CPL}} \text{token}$ ,  $\text{token}$  is what is called an *implicit static law* [23] of  $\mathcal{T}$ .<sup>9</sup>

<sup>9</sup>Implicit static laws are very closely related to veridical paradoxes [44]. It turns out

Modular theories have interesting properties. For example, consistency can be checked by just checking consistency of the static laws in  $\mathcal{S}$ : if  $\mathcal{T}$  is modular, then  $\mathcal{T} \models_{\overline{\text{PDL}}} \perp$  if and only if  $\mathcal{S} \models_{\overline{\text{CPL}}} \perp$ . Deduction of effect laws does not need the executability ones and vice versa. Deduction of an effect of a sequence of actions  $a_1; \dots; a_n$  (prediction) does not need to take into account the effect laws for actions other than  $a_1, \dots, a_n$ . This applies in particular to plan validation when deciding whether  $\langle a_1; \dots; a_n \rangle \varphi$  is the case.

Similar notions to modularity have been investigated in the literature on regulation consistency [7], Situation Calculus [2, 24], DL ontologies [8, 26] and also in dynamic logic [56]. For more details on modularity in action theories, see the work by Varzinczak [51].

**Theorem 5.1**  *$\mathcal{T}$  is modular if and only if the big model of  $\mathcal{T}$  is a model of  $\mathcal{T}$ .*

**Proof:** Let  $\mathcal{M}_{big} = \langle W_{big}, R_{big} \rangle$  be the big model of  $\mathcal{T}$ .

( $\Rightarrow$ ): By definition,  $\mathcal{M}_{big}$  is such that  $\models_{\mathcal{M}_{big}} \mathcal{S} \wedge \mathcal{E}$ . It remains to show that  $\models_{\mathcal{M}_{big}} \mathcal{X}$ . Let  $\varphi_i \rightarrow \langle a \rangle \top \in \mathcal{X}_a$ , and let  $w \in W_{big}$  be such that  $\models_w^{\mathcal{M}_{big}} \varphi_i$ . Therefore for all  $\varphi_j \in \mathfrak{Fml}$  such that  $\mathcal{T} \models_{\overline{\text{PDL}}} \varphi_j \rightarrow [a] \perp$ , we must have  $\not\models_w^{\mathcal{M}_{big}} \varphi_j$ , because  $\mathcal{T} \models_{\overline{\text{PDL}}} \neg(\varphi_i \wedge \varphi_j)$ , and as  $\mathcal{T}$  is modular,  $\mathcal{S} \models_{\overline{\text{CPL}}} \neg(\varphi_i \wedge \varphi_j)$ , and hence  $\models_{\mathcal{M}_{big}} \neg(\varphi_i \wedge \varphi_j)$ . Then by the construction of  $\mathcal{M}_{big}$ , there is some  $w' \in W_{big}$  such that  $\models_{w'}^{\mathcal{M}_{big}} \psi$  for all  $\varphi \rightarrow [a] \psi \in \mathcal{E}_a$  such that  $\models_w^{\mathcal{M}_{big}} \varphi$ . Thus  $R_a(w) \neq \emptyset$  and  $\models_{\mathcal{M}_{big}} \varphi_i \rightarrow \langle a \rangle \top$ .

( $\Leftarrow$ ): Suppose  $\mathcal{T}$  is not modular. Then there must be some  $\varphi \in \mathfrak{Fml}$  such that  $\mathcal{T} \models_{\overline{\text{PDL}}} \varphi$  and  $\mathcal{S} \not\models_{\overline{\text{CPL}}} \varphi$ . This means that there is  $v \in \text{val}(\mathcal{S})$  such that  $v \not\models \varphi$ . As  $v \in W_{big}$  (because  $W_{big}$  contains all possible valuations of  $\mathcal{S}$ ),  $\mathcal{M}_{big}$  is not a model of  $\mathcal{T}$ . ■

### 5.3 Correctness Under Modularity

The following theorem establishes that the semantic contraction of a formula  $\Phi$  from the set of models of an action theory  $\mathcal{T}$  produces models of some contracted theory in  $\mathcal{T}_{\Phi}^-$ .

**Theorem 5.2** *Let  $\mathcal{T}$  be modular, and  $\Phi$  be a law. For all  $\mathcal{M}' \in \mathcal{M}_{\Phi}^-$  such that  $\models_{\mathcal{M}'} \mathcal{T}$  for every  $\mathcal{M} \in \mathcal{M}$ , there is  $\mathcal{T}' \in \mathcal{T}_{\Phi}^-$  such that  $\models_{\mathcal{M}'}^{\mathcal{M}'} \mathcal{T}'$  for every  $\mathcal{M}' \in \mathcal{M}'$ .*

---

that they are not always intuitive. For a deep discussion on implicit static laws, see the article by Herzig and Varzinczak [27].

**Proof:** See Appendix A. ■

The next theorem establishes the other way round: models of theories in  $\mathcal{T}_{\Phi}^-$  are all models of the semantic contraction of  $\Phi$  from models of  $\mathcal{T}$ .

**Theorem 5.3** *Let  $\mathcal{T}$  be modular,  $\Phi$  a law, and  $\mathcal{T}' \in \mathcal{T}_{\Phi}^-$ . For all  $\mathcal{M}'$  such that  $\models^{\mathcal{M}'} \mathcal{T}'$ , there is  $\mathcal{M}' \in \mathcal{M}_{\Phi}^-$  such that  $\mathcal{M}' \in \mathcal{M}'$  and  $\models^{\mathcal{M}} \mathcal{T}$  for every  $\mathcal{M} \in \mathcal{M}$ .*

**Proof:** See Appendix B. ■

With these two theorems one gets correctness of the operators:

**Corollary 5.1** *Let  $\mathcal{T}$  be modular,  $\Phi$  a law, and  $\mathcal{T}' \in \mathcal{T}_{\Phi}^-$ . Then  $\mathcal{T}' \models_{\text{PDL}} \Psi$  if and only if  $\models^{\mathcal{M}'} \Psi$  for every  $\mathcal{M}' \in \mathcal{M}'$  such that  $\mathcal{M}' \in \mathcal{M}_{\Phi}^-$  for some  $\mathcal{M}$  such that  $\models^{\mathcal{M}} \mathcal{T}$  for all  $\mathcal{M} \in \mathcal{M}$ .*

**Proof:**

( $\Rightarrow$ ): Let  $\mathcal{M}'$  be such that  $\models^{\mathcal{M}'} \mathcal{T}'$ . By Theorem 5.3, there is  $\mathcal{M}' \in \mathcal{M}_{\Phi}^-$  such that  $\mathcal{M}' \in \mathcal{M}'$  for some  $\mathcal{M}$  such that  $\models^{\mathcal{M}} \mathcal{T}$  for all  $\mathcal{M} \in \mathcal{M}$ . From this and  $\mathcal{T}' \models_{\text{PDL}} \Psi$ , we have  $\models^{\mathcal{M}'} \Psi$ .

( $\Leftarrow$ ): Suppose  $\mathcal{T}' \not\models_{\text{PDL}} \Psi$ . (We show that there is some model  $\mathcal{M}' \in \mathcal{M}'$  such that  $\mathcal{M}' \in \mathcal{M}_{\Phi}^-$  for some  $\mathcal{M}$  with  $\models^{\mathcal{M}} \mathcal{T}$  for all  $\mathcal{M} \in \mathcal{M}$ , and  $\not\models^{\mathcal{M}'} \Psi$ .)

Given that  $\mathcal{T}$  is modular, by Lemma B.1  $\mathcal{T}'$  is modular, too. Then, by Lemma B.3, there is  $\mathcal{M}' = \langle \text{val}(S'), R' \rangle$  such that  $\not\models^{\mathcal{M}'} \Psi$ . Clearly  $\models^{\mathcal{M}'} \mathcal{T}'$ , and from Lemma B.4 the result follows. ■

## 6 Assessment of Postulates for Change

Do our action theory change operators satisfy the classical postulates for change? Before answering this question, one should ask: do our operators behave like revision or update operators? We here address this issue and then show which postulates for theory change are satisfied by our definitions.

### 6.1 Contraction or Erasure?

The distinction between revision/contraction and update/erasure for classical theories is historically controversial in the literature. The same is true for the case of modal theories describing actions and their effects. We here



rephrase Katsuno and Mendelzon's definitions [31] in our terms so that we can see to which one our method is closer.

In Katsuno and Mendelzon's view, contracting a law  $\Phi$  from an action theory  $\mathcal{T}$  intuitively means that the description of the possible behavior of the dynamic world  $\mathcal{T}$  must be adjusted to the possibility of  $\Phi$  being false. This amounts to selecting from the models of  $\neg\Phi$  those that are closest to models of  $\mathcal{T}$  and allow them as models of the result.

In contrast, update methods select, for each model  $\mathcal{M}$  of  $\mathcal{T}$ , the set of models of  $\Phi$  that are closest to  $\mathcal{M}$ . Erasing  $\Phi$  from  $\mathcal{T}$  means adding models to  $\mathcal{T}$ ; for each model  $\mathcal{M}$ , we add all those models closest to  $\mathcal{M}$  in which  $\Phi$  is false. Hence, from our constructions so far it seems that our operators are closer to update than to revision.

Moreover, according to Katsuno and Mendelzon's view [31], our change operators would also be classified as update because we make modifications in each model independently, i.e., without changing other models.<sup>10</sup> Besides that, in our setting a different ordering on the resulting models is induced by each model of  $\mathcal{T}$  (see Definitions 3.3, 3.7 and 3.10), which according to Katsuno and Mendelzon is a typical property of an update/erasure method.

Nevertheless, things get quite different when it comes to the postulates for theory change.

## 6.2 The Postulates

We here analyze the behavior of our action theory change operators w.r.t. Katsuno and Mendelzon's postulates and variants. Let  $\mathcal{T} = \mathcal{S} \cup \mathcal{E} \cup \mathcal{X}$  denote an action theory and  $\Phi$  denote a law.

**Monotonicity Postulate:**  $\mathcal{T} \models_{\text{PDL}} \mathcal{T}'$ , for all  $\mathcal{T}' \in \mathcal{T}_{\Phi}^{-}$ .

This postulate is our version of Katsuno and Mendelzon's (C1) and (E1) postulates for contraction and erasure, respectively, and is satisfied by our change operators. The proof is in Lemma A.1. Such a postulate is not satisfied by the operators proposed by Herzig *et al.* [21]: there when removing e.g. an executability law  $\varphi \rightarrow \langle a \rangle \top$  one may make  $\varphi \rightarrow [a] \perp$  valid in all models of the resulting theory.

**Preservation Postulate:** If  $\mathcal{T} \not\models_{\text{PDL}} \Phi$ , then  $\models_{\text{PDL}} \mathcal{T} \leftrightarrow \mathcal{T}'$ , for all  $\mathcal{T}' \in \mathcal{T}_{\Phi}^{-}$ .

---

<sup>10</sup>Even if when contracting an effect law from one particular model we need to check the other models of the theory, those are not modified.

This is Katsuno and Mendelzon's (C2) postulate. Our operators satisfy it as far as whenever  $\mathcal{T} \not\models_{\overline{\text{PDL}}} \Phi$ , then the models of the resulting theory are exactly the models of  $\mathcal{T}$ , because these are the minimal models falsifying  $\Phi$ .

The corresponding version of Katsuno and Mendelzon's (E2) postulate about erasure, i.e., if  $\mathcal{T} \models_{\overline{\text{PDL}}} \neg\Phi$ , then  $\models_{\overline{\text{PDL}}} \mathcal{T} \leftrightarrow \mathcal{T}'$ , for all  $\mathcal{T}' \in \mathcal{T}_{\Phi}^-$ , is clearly also satisfied by our operators as a special case of the postulate above. Satisfaction of (C2) indicates that our operators are closer to contraction than to erasure.

**Success Postulate:** If  $\mathcal{T} \not\models_{\overline{\text{PDL}}} \perp$  and  $\models_{\overline{\text{PDL}}} \Phi$ , then  $\mathcal{T}' \not\models_{\overline{\text{PDL}}} \Phi$ , for all  $\mathcal{T}' \in \mathcal{T}_{\Phi}^-$ .

This postulate is our version of Katsuno and Mendelzon's (C3) and (E3) postulates. If  $\Phi$  is a propositional  $\varphi \in \mathfrak{Fml}$ , our operators satisfy it, as long as the classical propositional change operator satisfies it. For the general case, however, as stated the postulate is not always satisfied. This is shown by the following example: let  $\mathcal{T} = \{\neg p, \langle a \rangle \top, p \rightarrow [a] \perp\}$ . Note that  $\mathcal{T}$  is modular and consistent. Now, contracting the (contingent) formula  $p \rightarrow \langle a \rangle \top$  from  $\mathcal{T}$  gives us  $\mathcal{T}' = \mathcal{T}$ . Clearly  $\mathcal{T}' \models_{\overline{\text{PDL}}} p \rightarrow \langle a \rangle \top$ . This happens because, despite not being a tautology,  $p \rightarrow \langle a \rangle \top$  is a 'trivial' formula w.r.t.  $\mathcal{T}$ : since  $\neg p$  is valid in all  $\mathcal{T}$ -models,  $p \rightarrow \langle a \rangle \top$  is trivially true in these models.

Fortunately, for all those formulas that are non-trivial consequences of the theory, our operators guarantee success of contraction:

**Theorem 6.1** *Let  $\mathcal{T}$  be consistent, and  $\Phi$  be an executability or an effect law such that  $\mathcal{S} \not\models_{\overline{\text{PDL}}} \Phi$ . If  $\mathcal{T}$  is modular, then  $\mathcal{T}' \not\models_{\overline{\text{PDL}}} \Phi$  for every  $\mathcal{T}' \in \mathcal{T}_{\Phi}^-$ .*

**Proof:** Suppose there is  $\mathcal{T}' \in \mathcal{T}_{\Phi}^-$  such that  $\mathcal{T}' \models_{\overline{\text{PDL}}} \Phi$ . As  $\mathcal{T}$  is modular, Corollary 5.1 gives us  $\models^{\mathcal{M}'} \Phi$  for every  $\mathcal{M}' \in \mathcal{M}'$  such that  $\mathcal{M}' \in \mathcal{M}_{\Phi}^-$ , where  $\mathcal{M} = \{\mathcal{M} : \models^{\mathcal{M}} \mathcal{T} \text{ and } \mathcal{M} = \langle \text{val}(\mathcal{S}), R \rangle\}$ .

If  $\models^{\mathcal{M}'} \Phi$  for every  $\mathcal{M}' \in \mathcal{M}'$ , then even for  $\mathcal{M}'' \in \mathcal{M}' \setminus \mathcal{M}$  we have  $\models^{\mathcal{M}''} \Phi$ . But  $\mathcal{M}'' \in \mathcal{M}_{\Phi}^-$  for some  $\mathcal{M} \in \mathcal{M}$ , and by definition  $\not\models^{\mathcal{M}''} \Phi$ . Hence  $\mathcal{M}_{\Phi}^- = \emptyset$ , and then the truth of  $\Phi$  in  $\mathcal{M}$  does not depend on  $R_a$ . Then, whether  $\Phi$  has the form  $\varphi \rightarrow \langle a \rangle \top$  or  $\varphi \rightarrow [a] \psi$ , for  $\varphi, \psi \in \mathfrak{Fml}$ , this holds only if  $\mathcal{S} \models_{\text{CPL}} \neg\varphi$  (see Definitions 3.1 and 3.5), in which case we get  $\mathcal{S} \models_{\overline{\text{PDL}}} \Phi$ . ■

**Equivalences Postulate:** If  $\models_{\overline{\text{PDL}}} \mathcal{T}_1 \leftrightarrow \mathcal{T}_2$  and  $\models_{\overline{\text{PDL}}} \Phi_1 \leftrightarrow \Phi_2$ , then  $\models_{\overline{\text{PDL}}} \mathcal{T}'_1 \leftrightarrow \mathcal{T}'_2$ , for  $\mathcal{T}'_1 \in (\mathcal{T}_1)_{\Phi_2}^-$  and  $\mathcal{T}'_2 \in (\mathcal{T}_2)_{\Phi_1}^-$ .

This postulate corresponds to Katsuno and Mendelzon's (C4) and (E4) postulates. Under modularity and the assumption that the propositional change operator satisfies (C4)/(E4), our operations satisfy this postulate:

**Theorem 6.2** *Let  $\mathcal{T}_1$  and  $\mathcal{T}_2$  be modular. If  $\models_{\text{PDL}} \mathcal{T}_1 \leftrightarrow \mathcal{T}_2$  and  $\models_{\text{PDL}} \Phi_1 \leftrightarrow \Phi_2$ , then for each  $\mathcal{T}'_1 \in (\mathcal{T}_1)_{\Phi_2}^-$  there is  $\mathcal{T}'_2 \in (\mathcal{T}_2)_{\Phi_1}^-$  such that  $\models_{\text{PDL}} \mathcal{T}'_1 \leftrightarrow \mathcal{T}'_2$ , and vice-versa.*

**Proof:** The proof follows straight from our results: since  $\models_{\text{PDL}} \mathcal{T}_1 \leftrightarrow \mathcal{T}_2$  and  $\models_{\text{PDL}} \Phi_1 \leftrightarrow \Phi_2$ , they have pairwise the same models. Hence, given  $\mathcal{M}$  such that  $\models^{\mathcal{M}} \mathcal{T}_1$  and  $\models^{\mathcal{M}} \mathcal{T}_2$ , the semantic contraction of  $\Phi_1$  and that of  $\Phi_2$  from  $\mathcal{M}$  have the same operations on  $\mathcal{M}$ . As  $\mathcal{T}_1$  and  $\mathcal{T}_2$  are modular, Corollary 5.1 guarantees we get the same syntactical results. Moreover, as the classical operator  $\ominus$  satisfies (C4)/(E4), it follows that  $\models_{\text{PDL}} \mathcal{T}'_1 \leftrightarrow \mathcal{T}'_2$ . ■

**Recovery Postulate:**  $\mathcal{T}' \cup \{\Phi\} \models_{\text{PDL}} \mathcal{T}$ , for all  $\mathcal{T}' \in \mathcal{T}_{\Phi}^-$ .

This is the action theory counterpart of Katsuno and Mendelzon's (C5) and (E5) postulates. Again we rely on modularity in order to satisfy it.

**Theorem 6.3** *Let  $\mathcal{T}$  be modular.  $\mathcal{T}' \cup \{\Phi\} \models_{\text{PDL}} \mathcal{T}$ , for all  $\mathcal{T}' \in \mathcal{T}_{\Phi}^-$ .*

**Proof:** If  $\mathcal{T} \not\models_{\text{PDL}} \Phi$ , because our operators satisfy the preservation postulate,  $\mathcal{T}' = \mathcal{T}$ , and then the result follows by monotonicity.

Let  $\mathcal{T} \models_{\text{PDL}} \Phi$ , and let  $\mathcal{M}'$  denote the set of all models of  $\mathcal{T}'$ . As  $\mathcal{T}$  is modular, by Corollary 5.1 every  $\mathcal{M}' \in \mathcal{M}'$  is such that either  $\models^{\mathcal{M}'} \mathcal{T}$  (and then  $\models^{\mathcal{M}'} \Phi$ ) or  $\mathcal{M}' \in \text{contract}(\mathcal{M}, \Phi)$  (and then  $\mathcal{M}' \in \mathcal{M}_{\Phi}^-$ ) for some  $\mathcal{M}$  such that  $\models^{\mathcal{M}} \mathcal{T}$ .

Let  $\mathcal{M}''$  denote the set of all models of  $\mathcal{T}' \cup \{\Phi\}$ . Clearly  $\mathcal{M}'' \subseteq \mathcal{M}'$ , by monotonicity. Moreover, every  $\mathcal{M}'' \in \mathcal{M}''$  is such that  $\models^{\mathcal{M}''} \Phi$ , hence  $\mathcal{M}'' \notin \mathcal{M}_{\Phi}^-$  for every  $\mathcal{M}$  such that  $\models^{\mathcal{M}} \mathcal{T}$ , and then  $\mathcal{M}'' \notin \text{contract}(\mathcal{M}, \Phi)$ , for any  $\mathcal{M}$  model of  $\mathcal{T}$ . Thus  $\mathcal{M}''$  is a model of  $\mathcal{T}$  and then  $\mathcal{T}' \cup \{\Phi\} \models_{\text{PDL}} \mathcal{T}$ . ■

Let  $\bigvee \mathcal{T}_{\Phi}^-$  denote the disjunction of all  $\mathcal{T}'$  in  $\mathcal{T}_{\Phi}^-$ .

**Disjunctive rule:**  $(\mathcal{T}_1 \vee \mathcal{T}_2)_{\Phi}^-$  is equivalent to  $\bigvee (\mathcal{T}_1)_{\Phi}^- \vee \bigvee (\mathcal{T}_2)_{\Phi}^-$ .

This is our version of (E8) erasure postulate by Katsuno and Mendelzon. Clearly our syntactical operators do not manage to contract a law from a disjunction of theories  $\mathcal{T}_1 \vee \mathcal{T}_2$ . Nevertheless, by proving that it holds in the semantics, from the correctness of our operators, we get an equivalent operation. Again the fact that the theories under concern are modular gives us the result.

**Theorem 6.4** *Let  $\mathcal{T}_1$  and  $\mathcal{T}_2$  be modular, and  $\Phi$  be a law. Then*

$$\models_{\text{PDL}} \bigvee (\mathcal{T}_1 \vee \mathcal{T}_2)_{\Phi}^- \leftrightarrow (\bigvee (\mathcal{T}_1)_{\Phi}^- \vee \bigvee (\mathcal{T}_2)_{\Phi}^-)$$

**Proof:**

( $\Leftarrow$ ): Let  $\mathcal{M}'$  be such that  $\models^{\mathcal{M}'} \bigvee (\mathcal{T}_1)_{\Phi}^- \vee \bigvee (\mathcal{T}_2)_{\Phi}^-$ . Then  $\models^{\mathcal{M}'} \bigvee (\mathcal{T}_1)_{\Phi}^-$  or  $\models^{\mathcal{M}'} \bigvee (\mathcal{T}_2)_{\Phi}^-$ . Suppose  $\models^{\mathcal{M}'} \bigvee (\mathcal{T}_1)_{\Phi}^-$  (the other case is analogous). Then there is  $(\mathcal{T}_1)' \in (\mathcal{T}_1)_{\Phi}^-$  such that  $\models^{\mathcal{M}'} (\mathcal{T}_1)'$ . Then by Corollary 5.1, there is  $\mathcal{M}' \in \mathcal{M}_{\Phi}^-$  such that  $\mathcal{M}' \in \mathcal{M}'$ , for  $\mathcal{M}$  a set of models of  $\mathcal{T}_1$ . Then  $\mathcal{M}'$  is a model resulting from contracting  $\Phi$  from models of  $\mathcal{T}_1$ , and then  $\mathcal{M}'$  also results from contracting  $\Phi$  in models of  $\mathcal{T}_1 \vee \mathcal{T}_2$ , viz. those models of  $\mathcal{T}_1$ . Then by Corollary 5.1, there is  $(\mathcal{T}_1 \vee \mathcal{T}_2)' \in (\mathcal{T}_1 \vee \mathcal{T}_2)_{\Phi}^-$  such that  $\models^{\mathcal{M}'} (\mathcal{T}_1 \vee \mathcal{T}_2)'$ , and then  $\models^{\mathcal{M}'} \bigvee (\mathcal{T}_1 \vee \mathcal{T}_2)_{\Phi}^-$ .

( $\Rightarrow$ ): Let  $\mathcal{M}'$  be such that  $\models^{\mathcal{M}'} \bigvee (\mathcal{T}_1 \vee \mathcal{T}_2)_{\Phi}^-$ . Then there is  $(\mathcal{T}_1 \vee \mathcal{T}_2)' \in (\mathcal{T}_1 \vee \mathcal{T}_2)_{\Phi}^-$  such that  $\models^{\mathcal{M}'} (\mathcal{T}_1 \vee \mathcal{T}_2)'$ . By Corollary 5.1, there is  $\mathcal{M}' \in \mathcal{M}_{\Phi}^-$  such that  $\mathcal{M}' \in \mathcal{M}'$ , for  $\mathcal{M}$  a set of models of  $\mathcal{T}_1 \vee \mathcal{T}_2$ . Then  $\mathcal{M}'$  is a model resulting from contracting  $\Phi$  from models of  $\mathcal{T}_1 \vee \mathcal{T}_2$ . Hence  $\mathcal{M}'$  results from contracting  $\Phi$  from models of  $\mathcal{T}_1$  or from models of  $\mathcal{T}_2$ . Suppose the former is the case (the second is analogous). Then by Corollary 5.1 there is  $(\mathcal{T}_1)' \in (\mathcal{T}_1)_{\Phi}^-$  such that  $\models^{\mathcal{M}'} (\mathcal{T}_1)'$ , and then  $\models^{\mathcal{M}'} \bigvee (\mathcal{T}_1)_{\Phi}^-$ . ■

We have thus shown that our constructions satisfy (E8) postulate. Nevertheless there is no evidence whether it is really expected here. This supports our position that our operators' behavior is closer to contraction than to erasure. To finish up we state a new postulate:

**Preservation of modularity:** If  $\mathcal{T}$  is modular, then every  $\mathcal{T}' \in \mathcal{T}_{\Phi}^-$  is modular.

Changing a modular theory should not make it nonmodular. This is not a standard postulate, but we think that as a good property modularity should be preserved across changing an action theory. If so, this means that whether a theory is modular or not can be checked once for all and one does not need to care about it during the future evolution of the action theory, i.e., when other changes will be made on it. Our operators satisfy this postulate and the proof is given in Appendix B.

## 7 A Semantics for Action Theory Revision

So far we have analyzed the case of contraction: when evolving a theory one realizes that it is too strong and hence it has to be weakened. Let's now take a look at the other way round, i.e., the theory is too liberal and the agent discovers new *laws* about the world that should be added to her beliefs, which amounts to strengthening them.

Suppose the action theory of our scenario example were initially stated as follows:

$$\mathcal{T} = \left\{ \begin{array}{l} \text{coffee} \rightarrow \text{hot}, \text{token} \rightarrow \langle \text{buy} \rangle \top, \\ \neg \text{coffee} \rightarrow [\text{buy}] \text{coffee}, \neg \text{token} \rightarrow [\text{buy}] \perp, \\ \text{coffee} \rightarrow [\text{buy}] \text{coffee}, \text{hot} \rightarrow [\text{buy}] \text{hot} \end{array} \right\}$$

Then the big-model of  $\mathcal{T}$  is as shown in Figure 12.

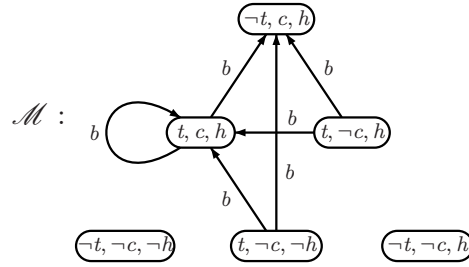


Figure 12: Model of the new initial action domain description.

Looking at model  $\mathcal{M}$  in Figure 12 we can see that, for example, the agent does not know that she loses her token every time she buys coffee at the machine. This is a new law that she should incorporate to her knowledge base at some stage of her action theory evolution.

Contrary to contraction, where we want the negation of some law to become *satisfiable*, in revision we want to make a new law *valid*. This means that one has to eliminate all cases satisfying its negation. This depicts the duality between revision and contraction: whereas in the latter one invalidates a formula by making its negation satisfiable, in the former one makes a formula valid by forcing its negation to be unsatisfiable prior to adding the new law to the theory.

The idea behind our semantics is as follows: we initially have a set of models  $\mathcal{M}$  in which a given formula  $\Phi$  is (potentially) not valid, i.e.,  $\Phi$  is

(possibly) not true in every model in  $\mathcal{M}$ . In the result we want to have only models of  $\Phi$ . Adding  $\Phi$ -models to  $\mathcal{M}$  is of no help. Moreover, adding models makes us lose laws: the corresponding resulting theory would be more liberal.

One solution amounts to deleting from  $\mathcal{M}$  those models that are not  $\Phi$ -models. Of course removing only some of them does not solve the problem, we must delete every such a model. By doing that, all resulting models will be models of  $\Phi$ . (This corresponds to *theory expansion*, when the resulting theory is satisfiable.) However, if  $\mathcal{M}$  contains no model of  $\Phi$ , we will end up with  $\emptyset$ . Consequence: the resulting theory is inconsistent. (This is the main revision problem.) In this case the solution is to *substitute* each model  $\mathcal{M}$  in  $\mathcal{M}$  by its *nearest modification*  $\mathcal{M}_\Phi^*$  that makes  $\Phi$  true. This lets us to keep as close as possible to the original models we had. But, what if for one model in  $\mathcal{M}$  there are several minimal (incomparable) modifications of it validating  $\Phi$ ? In that case we will consider all of them. The result will also be a *list of models*  $\mathcal{M}_\Phi^*$ , all being models of  $\Phi$ .

Before defining revision of sets of models, we present what modifications of (individual) models are.

## 7.1 Revising a Model by a Static Law

Suppose that our coffee deliverer agent discovers that the only hot beverage that is served on the machine is coffee. In this case, she might want to revise her beliefs with the new static law  $\neg coffee \rightarrow \neg hot$ : she cannot hold a hot beverage that is not a coffee.

Considering the model depicted in Figure 12, one sees that the formula  $\neg coffee \wedge hot$  is satisfiable. As we do not want this, the first step is to *remove* all worlds in which  $\neg coffee \wedge hot$  is true. The second step is to guarantee that all the remaining worlds satisfy the new law. Such an issue has been largely addressed in the literature on propositional belief base revision and update [15, 55, 31, 22]. Here we can achieve that with a semantics similar to that of classical revision operators: basically one can change the set of possible valuations, by removing or adding worlds.

In our example, removing the possible worlds  $\{t, \neg c, h\}$  and  $\{\neg t, \neg c, h\}$  would do the job (there is no need to add new valuations since the new incoming law is satisfied in at least one world of the resulting model).

The delicate point in removing worlds is that this may have as consequence the loss of some executability laws: in the example, if there were some arrow pointing from some world  $w$  to say  $\{\neg t, \neg c, h\}$ , then removing the latter from the model would make the action under concern no longer

executable in  $w$ , if it was the only arrow labeled by that action leaving it. From a semantic point of view, this is intuitive: if the state of the world to which we could move is no longer possible, then we do not have a transition to that state anymore. Hence, if that transition was the only one we had, it is natural to lose it.

Similarly, one could ask what to do with the accessibility relation if new worlds are added, i.e., when expansion is not possible. Following the discussion in Section 3.3, we here prefer not to systematically add new arrows to the accessibility relation, and postpone correction of executability laws, if needed. This approach is debatable, but with the information we have at hand, this is the safest way of changing static laws.

The semantics for revision of one model by a static law is as follows:

**Definition 7.1** Let  $\mathcal{M} = \langle W, R \rangle$ .  $\mathcal{M}' = \langle W', R' \rangle \in \mathcal{M}_\varphi^*$  if and only if:

- $W' = (W \setminus \text{val}(\neg\varphi)) \cup \text{val}(\varphi)$
- $R' \subseteq R$

Clearly  $\models^{\mathcal{M}'} \varphi$  for each  $\mathcal{M}' \in \mathcal{M}_\varphi^*$ . The minimal models resulting from revising a model  $\mathcal{M}$  by  $\varphi$  are those closest to  $\mathcal{M}$  w.r.t.  $\preceq_{\mathcal{M}}$ :

**Definition 7.2** Let  $\mathcal{M}$  be a model and  $\varphi$  a static law.  $\text{revise}(\mathcal{M}, \varphi) = \bigcup \min\{\mathcal{M}_\varphi^*, \preceq_{\mathcal{M}}\}$ .

In the example of model  $\mathcal{M}$  in Figure 12,  $\text{revise}(\mathcal{M}, \neg\text{coffee} \rightarrow \neg\text{hot})$  is the singleton  $\{\mathcal{M}'\}$ , where  $\mathcal{M}'$  is as shown in Figure 13.

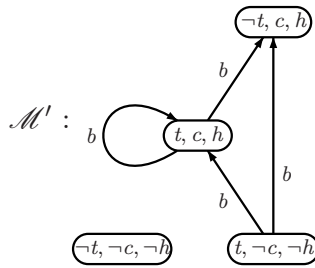


Figure 13: Model resulting from revising the model  $\mathcal{M}$  in Figure 12 with  $\neg\text{coffee} \rightarrow \neg\text{hot}$ .

## 7.2 Revising a Model by an Effect Law

Let's suppose now that our agent eventually discovers that after buying coffee she does not keep her token. This means that her theory should now be revised by the new effect law  $token \rightarrow [buy]\neg token$ . Looking at model  $\mathcal{M}$  in Figure 12, this amounts to guaranteeing that the formula  $token \wedge \langle buy \rangle token$  is satisfiable in none of its worlds. To do that, we have to look at all the worlds satisfying this formula (if any) and

- either make  $token$  false in each of these worlds,
- or make  $\langle buy \rangle token$  false in all of them.

If we chose the first option, we will essentially flip the truth value of literal  $token$  in the respective worlds, which changes the set of valuations of the model. If we chose the latter, we will basically remove  $buy$ -arrows leading to  $token$ -worlds. In that case, a change in the accessibility relation will be made.

In our example, we have that the possible worlds  $\{token, coffee, hot\}$ ,  $\{token, \neg coffee, hot\}$  and  $\{token, \neg coffee, \neg hot\}$  satisfy  $token \wedge \langle buy \rangle token$  and all they have to change.

Flipping  $token$  in all these worlds to  $\neg token$  would do the job, but would also have as consequence the introduction of a new static law:  $\neg token$  would now be valid, i.e., the agent never has a token.

Here we think that changing action laws should not have as side effect a change in the static laws. Given their special status, these should change only if explicitly required (see above). In this case, each world satisfying  $token \wedge \langle buy \rangle token$  has to be changed so that  $\langle buy \rangle token$  is no longer true in it. In our example, we should remove the arrows  $(\{token, coffee, hot\}, \{token, coffee, hot\})$ ,  $(\{token, \neg coffee, hot\}, \{token, coffee, hot\})$  and  $(\{token, \neg coffee, \neg hot\}, \{token, coffee, hot\})$ .

The semantics of one model revision for the case of a new effect law is:

**Definition 7.3** Let  $\mathcal{M} = \langle W, R \rangle$ .  $\mathcal{M}' = \langle W', R' \rangle \in \mathcal{M}_{\varphi \rightarrow [a]\psi}^*$  if and only if:

- $W' = W$
- $R' \subseteq R$
- If  $(w, w') \in R \setminus R'$ , then  $\models_w^{\mathcal{M}} \varphi$
- $\models^{\mathcal{M}'} \varphi \rightarrow [a]\psi$



The minimal models resulting from the revision of a model  $\mathcal{M}$  by a new effect law are those that are closest to  $\mathcal{M}$  w.r.t.  $\preceq_{\mathcal{M}}$ :

**Definition 7.4** *Let  $\mathcal{M}$  be a model and  $\varphi \rightarrow [a]\psi$  an effect law.  $\text{revise}(\mathcal{M}, \varphi \rightarrow [a]\psi) = \bigcup \min\{\mathcal{M}_{\varphi \rightarrow [a]\psi}^*, \preceq_{\mathcal{M}}\}$ .*

Taking again  $\mathcal{M}$  as in Figure 12,  $\text{revise}(\mathcal{M}, \text{token} \rightarrow [\text{buy}]\neg\text{token})$  will be the singleton  $\{\mathcal{M}'\}$  (Figure 14).

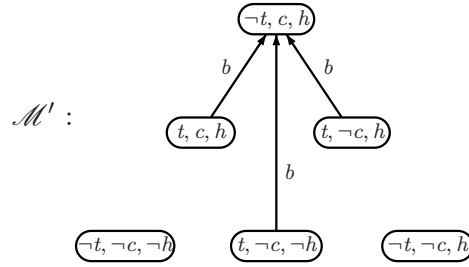


Figure 14: Model resulting from revising the model  $\mathcal{M}$  in Figure 12 with the new effect law  $\text{token} \rightarrow [\text{buy}]\neg\text{token}$ .

### 7.3 Revising a Model by an Executability Law

Let us now suppose that in some stage it has been decided to grant free coffee to everybody. Faced with this information, the agent will now revise her laws to reflect the fact that *buy* can also be executed in  $\neg\text{token}$ -contexts:  $\neg\text{token} \rightarrow \langle\text{buy}\rangle\top$  is a new executability law (and hence we will have  $\langle\text{buy}\rangle\top$  in all new models of the agent's beliefs).

Considering again the model in Figure 12, we observe that  $\neg(\neg\text{token} \rightarrow \langle\text{buy}\rangle\top)$  is satisfiable in  $\mathcal{M}$ . Hence we must throw  $\neg\text{token} \wedge [\text{buy}]\perp$  away to ensure the new formula becomes true.

To remove  $\neg\text{token} \wedge [\text{buy}]\perp$  we have to look at all worlds satisfying it and modify  $\mathcal{M}$  so that they no longer satisfy that formula. Given worlds  $\{\neg\text{token}, \neg\text{coffee}, \neg\text{hot}\}$  and  $\{\neg\text{token}, \neg\text{coffee}, \text{hot}\}$ , we have two options: change the interpretation of *token* or add new arrows leaving these worlds. A question that arises is ‘what choice is more drastic: change a world or an arrow’? Again, here we think that changing the world’s content (the valuation) is more drastic, as the existence of such a world was foreseen by some static law and is hence assumed to be as it is, unless we have enough information supporting the contrary, in which case we explicitly change the

static laws (see above). Thus we shall add a new *buy*-arrow from each of  $\{\neg token, \neg coffee, \neg hot\}$  and  $\{\neg token, \neg coffee, hot\}$ .

Having agreed on that, the issue now is: which worlds should the new arrows point to? Recalling the reasoning developed in Section 3.2, in order to comply with minimal change, the new arrows shall point to worlds that are relevant targets of each of the  $\neg token$ -worlds in question. In our example,  $\{\neg token, coffee, hot\}$  is the only relevant target world here: the two other  $\neg token$ -worlds violate the effect *coffee* of *buy*, while the three *token*-worlds would make us violate the frame axiom  $\neg token \rightarrow [buy]\neg token$ .

The semantics for one model revision by a new executability law is as follows:

**Definition 7.5** *Let  $\mathcal{M} = \langle W, R \rangle$ .  $\mathcal{M}' = \langle W', R' \rangle \in \mathcal{M}_{\varphi \rightarrow \langle a \rangle \top}^*$  if and only if:*

- $W' = W$
- $R \subseteq R'$
- *If  $(w, w') \in R' \setminus R$ , then  $w' \in RelTarget(w, \varphi \rightarrow [a]\perp, \mathcal{M}, \mathcal{M})$*
- $\models^{\mathcal{M}'} \varphi \rightarrow \langle a \rangle \top$

The minimal models resulting from revising a model  $\mathcal{M}$  by a new executability law are those closest to  $\mathcal{M}$  w.r.t.  $\preceq_{\mathcal{M}}$ :

**Definition 7.6** *Let  $\mathcal{M}$  be a model and  $\varphi \rightarrow \langle a \rangle \top$  be an executability law.  $revise(\mathcal{M}, \varphi \rightarrow \langle a \rangle \top) = \bigcup \min\{\mathcal{M}_{\varphi \rightarrow \langle a \rangle \top}^*, \preceq_{\mathcal{M}}\}$ .*

In our running example,  $revise(\mathcal{M}, \neg token \rightarrow \langle buy \rangle \top)$  is the singleton  $\{\mathcal{M}'\}$ , where  $\mathcal{M}'$  is as shown in Figure 15.

## 7.4 Revising Sets of Models

Up until now we have seen what the revision of single models means. This is needed when expansion by the new law is not possible due to inconsistency. We here give a unified definition of revision of a set of models  $\mathcal{M}$  by a new law  $\Phi$ :

**Definition 7.7** *Let  $\mathcal{M}$  be a set of models and  $\Phi$  a law. Then*

$$\mathcal{M}_{\Phi}^* = \begin{cases} \mathcal{M} \setminus \{\mathcal{M} : \not\models^{\mathcal{M}} \Phi\}, & \text{if there is } \mathcal{M} \in \mathcal{M} \text{ s.t. } \models^{\mathcal{M}} \Phi \\ \bigcup_{\mathcal{M} \in \mathcal{M}} revise(\mathcal{M}, \Phi), & \text{otherwise} \end{cases}$$

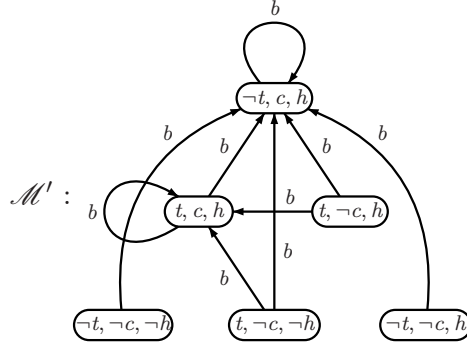


Figure 15: The result of revising model  $\mathcal{M}$  in Figure 12 by the new executability law  $\neg token \rightarrow \langle buy \rangle \top$ .

Observe that Definition 7.7 comprises both *expansion* and *revision*: in the first one, simple addition of the new law gives a satisfiable theory; in the latter a deeper change is needed to get rid of inconsistency.

## 8 Related Work

To the best of our knowledge, the first work on updating an action domain description is that by Li and Pereira [33] in a narrative-based action description language [16]. Contrary to us, however, they mainly investigate the problem of updating the narrative with new observed *facts* and (possibly) with occurrences of actions that explain those facts. This amounts to updating a given state/configuration of the world (in our terms, what is true in a possible world) and focusing on the models of the narrative in which some actions took place (in our terms, the models of the action theory with a particular sequence of action executions). Clearly the models of the action laws remain the same.

Baral and Lobo [4] introduce extensions of action languages that allow for some causal laws to be stated as defeasible. Their work is similar to ours in that they also allow for weakening of laws: in their setting, effect propositions can be replaced by what they call defeasible (weakened versions of) effect propositions. Our approach is different from theirs in the way executability laws are dealt with. Here executability laws are explicit and we are also able to contract them. This feature is important when the qualification problem [37] is considered: we may always discover contexts that preclude the execution of a given action (cf. the Introduction).

Liberatore [34] proposes a framework for reasoning about actions in

which it is possible to express a given semantics of belief update, like Winslett’s [55] and Katsuno and Mendelzon’s [31]. This means it is the formalism, essentially an action description language, that is used to describe updates (the change of propositions from one state of the world to another) by expressing them as laws in the action theory.

The main difference between Liberatore’s work and Li and Pereira’s is that, despite not being concerned, at least a priori, with changing action laws, Liberatore’s framework allows for abductively introducing in the action theory new effect propositions (effect laws, in our terms) that consistently explain the occurrence of an event.

The work by Eiter *et al.* [12, 13] is similar to ours in that they also propose a framework that is oriented to updating action laws. They mainly investigate the case where e.g. a new effect law is added to the description (and then has to be true in all models of the modified theory). This problem is the dual of contraction and is then closer to our definition of revision (cf. Section 7).

In Eiter *et al.*’s framework, action theories are described in a variant of a narrative-based action description language. Like in the present work, the semantics is also in terms of transition systems: directed graphs having arrows (action occurrences) linking nodes (configurations of the world). Contrary to us, however, the minimality condition on the outcome of the update is in terms of inclusion of sets of laws, which means the approach is more syntax oriented.

In their setting, during an update an action theory  $\mathcal{T}$  is seen as composed of two pieces,  $\mathcal{T}_u$  and  $\mathcal{T}_m$ , where  $\mathcal{T}_u$  stands for the part of  $\mathcal{T}$  that is not supposed to change and  $\mathcal{T}_m$  contains the laws that may be modified. In our terms, when contracting a static law we would have  $\mathcal{T}_m = \mathcal{S} \cup \mathcal{X}_a$ , when contracting an executability  $\mathcal{T}_m = \mathcal{X}_a$ , and when contracting effects laws  $\mathcal{T}_m = \mathcal{E}_a^-$ . The difference here is that in our approach it is always clear what laws should not change in a given type of contraction, and  $\mathcal{T}_u$  and  $\mathcal{T}_m$  do not need to be explicitly specified prior to the update.

Their approach and ours can both be described as *constraint-based* update, in that the theory change is carried out relative to some restrictions (a set of laws that we want to hold in the result). In our framework, for example, all changes in the action laws are relative to the set of static laws  $\mathcal{S}$  (and that is why we concentrate on models of  $\mathcal{T}$  having  $val(\mathcal{S})$  as worlds). When changing a law, we want to keep the same set of states. The difference w.r.t. Eiter *et al.*’s approach is that there it is also possible to update a theory relatively to e.g. executability laws: when expanding  $\mathcal{T}$  with a new

effect law, one may want to constrain the change so that the action under concern is guaranteed to be executable in the result.<sup>11</sup> As shown in the referred work, this may require the withdrawal of some static law. Hence, in Eiter *et al.*'s framework, static laws do not have the same status as in ours.

Herzig *et al.* [21] define a method for action theory contraction that, despite the similarity with the current work and the common underlying motivations, is more limited than the present constructions.

First, with the referred approach we do not get minimal change. For example, in the referred work the operator for contracting executability laws is such that in the resulting theory the modified set of executabilities is given by

$$\mathcal{X}_a^- = \{(\varphi_i \wedge \neg\varphi) \rightarrow \langle a \rangle \top : \varphi_i \rightarrow \langle a \rangle \top \in \mathcal{X}_a\}$$

which, according to its semantics, gives theories among whose models are those resulting from removing arrows from *all*  $\varphi$ -worlds. A similar comment can be made w.r.t. contraction of effect laws.

Second, Herzig *et al.*'s contraction method does not satisfy most of the postulates for theory change that we have addressed in Section 6. Besides not satisfying the monotonicity postulate, it does not satisfy the preservation one. To witness, suppose we have a language with only one atom  $p$ , and the model  $\mathcal{M}$  depicted in Figure 16.



Figure 16: Counter-example to preservation in the method of contraction by Herzig *et al.* [21].

Then  $\models^{\mathcal{M}} p \rightarrow [a]\neg p$  and  $\not\models^{\mathcal{M}} [a]\neg p$ . Now the contraction operator defined there is such that when removing  $[a]\neg p$  from  $\mathcal{M}$  yields the model  $\mathcal{M}'$  in Figure 16 such that  $R'_a = W \times W$ . Then  $\not\models^{\mathcal{M}'} p \rightarrow [a]\neg p$ , i.e., the effect law  $p \rightarrow [a]\neg p$  is not preserved.

---

<sup>11</sup>We could simulate that in our approach with two successive modifications of  $\mathcal{T}$ : first adding the effect law and then an executability law (cf. Section 7).

## 9 Comments

In this section we make some comments about possible modifications or improvements in our constructions so far.

### 9.1 Other Distance Notions

Here we have used a model distance based on symmetric differences between sets. This distance is quite close to Winslett's [55] notion of closeness between interpretations in the Possible Models Approach (PMA). Instead of it, however, we could have considered other distance notions as well, like e.g. Dalal's [9] distance, Hamming distance [18], or weighted distance. Due to space limitations, we do not develop a thorough comparison among all these distances here. We nevertheless do show that with a cardinality-based distance, for example, we may not always get the intended result.

Let  $\text{card}(X)$  denote the number of elements in set  $X$ . Then suppose our closeness between PDL-models was defined as follows:

**Definition 9.1 (Cardinality-based closeness between PDL-Models)**

Let  $\mathcal{M} = \langle W, R \rangle$  be a model. Then  $\mathcal{M}' = \langle W', R' \rangle$  is at least as close to  $\mathcal{M}$  as  $\mathcal{M}'' = \langle W'', R'' \rangle$ , noted  $\mathcal{M}' \leq_{\mathcal{M}} \mathcal{M}''$ , if and only if

- either  $\text{card}(W \dot{-} W') \leq \text{card}(W \dot{-} W'')$
- or  $\text{card}(W \dot{-} W') = \text{card}(W \dot{-} W'')$  and  $\text{card}(R \dot{-} R') \leq \text{card}(R \dot{-} R'')$

Such a notion of distance is closely related to Dalal's [9] closeness.

Since when contracting a static law  $\varphi$  from a model  $\mathcal{M}$  we usually add one new possible world, it is easy to see that with this cardinality-based distance we get the same result in  $\text{contract}(\mathcal{M}, \varphi)$  as with the distance from Definition 2.10.

When it comes to contraction of action laws, and then changing the accessibility relations, however, this cardinality-based distance does not seem to fit with the intuitions. To witness, consider the model  $\mathcal{M}$  in Figure 17, which satisfies the executability law  $p_1 \rightarrow \langle a \rangle \top$ .

Then,  $\mathcal{M}_{p_1 \rightarrow \langle a \rangle \top}^- = \{\mathcal{M}', \mathcal{M}''\}$ , where  $\mathcal{M}'$  and  $\mathcal{M}''$  are as depicted in Figure 18.

Note that  $\mathcal{M}''$  is an intended contracted model. However, with the cardinality-based distance above we will get  $\{\mathcal{M}\}_{p_1 \rightarrow \langle a \rangle \top}^- = \{\{\mathcal{M}, \mathcal{M}'\}\}$ . We do not have  $\{\mathcal{M}, \mathcal{M}''\}$  in the result since  $\mathcal{M}' \leq_{\mathcal{M}} \mathcal{M}''$ : in  $\mathcal{M}'$  only one arrow has been removed, while in  $\mathcal{M}''$  two.

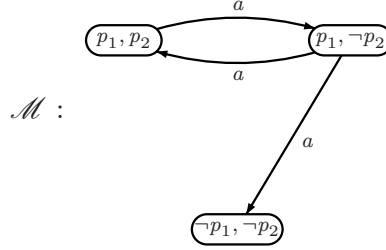


Figure 17: A model  $\mathcal{M}$  satisfying  $p_1 \rightarrow \langle a \rangle \top$ .

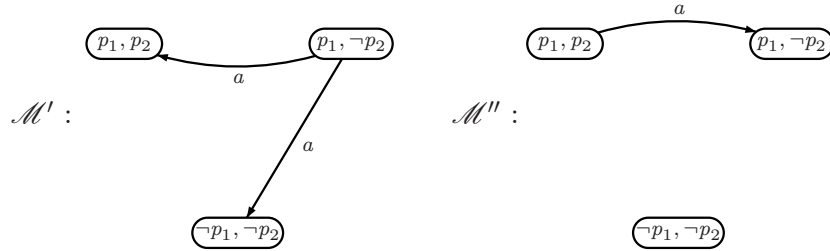


Figure 18: Models resulting from contracting  $p_1 \rightarrow \langle a \rangle \top$  in the model  $\mathcal{M}$  of Figure 17.

## 9.2 Inducing Executability

Regarding the semantics for contracting static laws, we could try to go further and at least make a guess about what executability laws we should preserve. Before doing that, we need a definition.

**Definition 9.2 (Closeness between Valuations)** *Let  $v$  be a propositional valuation. The valuation  $v'$  is as close to  $v$  as  $v''$ , noted  $v' \preceq_v v''$ , if and only if  $v \dot{-} v' \subseteq v \dot{-} v''$ .*

So the distance between valuations  $v_1$  and  $v_2$  is the set of literals on which they differ:  $v_1 \dot{-} v_2 = \{\ell : v_1 \models \ell \text{ and } v_2 \not\models \ell\} \cup \{\ell : v_2 \models \ell \text{ and } v_1 \not\models \ell\}$ .

Our argument now is as follows: when adding a new world, we can look at its contents and see what happens in worlds that are similar to it (by

similar here we mean the possible worlds that are closest to it). A priori and intuitively we can expect that if we put a new arrow leaving the new world, it will neither point to a world that is the target of no other world, nor point to a world that is not closest to it. It is reasonable to expect that in the new world a given action may have a behavior that is quite similar to that which it has in the worlds that are closest to the new one. Hence we select the worlds whose distance to the new one is minimal, look at where the arrows leaving them point to, and then point the new arrow there. With a similar argument, we can decide which arrows targeting the new world add to the model. The definition below formalizes this.

**Definition 9.3** *Let  $\mathcal{M} = \langle W, R \rangle$ .  $\mathcal{M}' = \langle W', R' \rangle \in \mathcal{M}_\varphi^-$  if and only if*

- $W \subseteq W'$
- $R \subseteq R'$
- *If  $w' \in W' \setminus W$ , then  $R'(w') \subseteq R(w)$  and  $R'^{-1}(w') \subseteq R^{-1}(w)$ , where  $w \in \bigcup \min\{W, \preceq_{w'}\}$*
- *There is  $w \in W'$  s.t.  $\not\models_w^{\mathcal{M}'} \varphi$*

With this new definition, what we do is suppose that some of the known laws for the other worlds can still be true in the new state, by analogy to the other possible states. In a similar way, when facing a new situation, we may wonder how we got there. Again, by analogy with known states, we could expect that we get to the new state coming from a state that usually produces something similar to what we have now in front of us. In this case we have a kind of abduction-like reasoning that may of course be wrong but that is not illegal.

Although intuitive, at least in its motivation, adopting Definition 9.3 could have some undesirable side effects. For example, if in the semantics we decide to add new arrows pointing from and to the new added world, then our corresponding operator may not satisfy the monotonicity postulate. To see, let

$$\mathcal{T} = \left\{ \begin{array}{l} p_1 \otimes p_2, p_1 \rightarrow [a]p_2, \\ p_1 \rightarrow \langle a \rangle \top, p_2 \rightarrow [a] \perp \end{array} \right\}$$

The only model of  $\mathcal{T}$  is  $\mathcal{M} = \langle W, R \rangle$  such that  $W = \{\{p_1, \neg p_2\}, \{\neg p_1, p_2\}\}$  and  $R = \{(\{p_1, \neg p_2\}, \{\neg p_1, p_2\})\}$  (Figure 19).



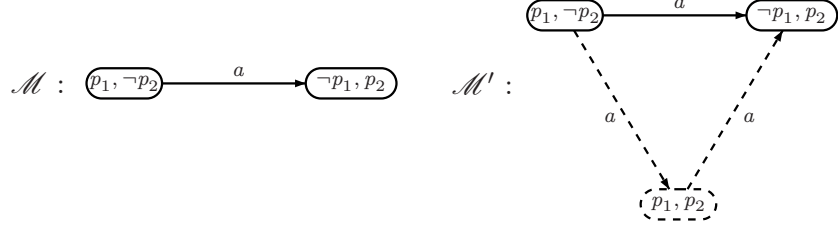


Figure 19: Counter-example to monotonicity when adding arrows to/from new worlds in the semantics of static law contraction.  $\mathcal{M}$  denotes the original model of  $\mathcal{T}$ , while  $\mathcal{M}'$  shows the new added world and the candidate arrows to add to  $R_a$ .

If we contract  $p_1 \rightarrow \neg p_2$  from  $\mathcal{T}$ , in the semantic result we have only the model  $\mathcal{M}'$  in Figure 19 such that  $\models^{\mathcal{M}'} p_1 \rightarrow \langle a \rangle \langle a \rangle p_2$ . Then, we would have  $\mathcal{T}' \models_{\text{PDL}} p_1 \rightarrow \langle a \rangle \langle a \rangle p_2$ , and then  $\mathcal{T} \not\models_{\text{PDL}} \mathcal{T}'$ .

The very issue with such a semantic characterization however would be how to capture it at the syntactic level: what syntax operator for change should we have in order to capture this closeness between possible worlds? More importantly, since we may be wrong about a guess regarding the executability or an effect of a given action, how can it be rolled back in the new theory? These are open questions that we leave for further investigation.

## 10 Concluding Remarks

In this work we have given a semantics for action theory change in terms of distances between models that captures the notion of minimal change. We have given algorithms to contract a formula from a theory that terminate and are correct w.r.t. the semantics (Corollary 5.1). We have shown the importance that modularity has in this result and in others.

Under modularity, our operators satisfy all the postulates for contraction. This supports the thesis that our modularity notion is fruitful.

By forcing formulas to be explicitly stated in their respective modules (and thus possibly making them inferable in independently different ways), modularity intuitively could be seen to diminish elaboration tolerance [38].

For instance, when contracting a Boolean formula  $\varphi$  in a non-modular theory, it seems reasonable to expect not to change the set of static laws  $\mathcal{S}$ , while the theory being modular surely forces changing such a module.

It is not difficult, however, to conceive non-modular theories in which contraction of a formula  $\varphi$  may demand a change in  $\mathcal{S}$  as well. As an example, suppose  $\mathcal{S} = \{\varphi_1 \rightarrow \varphi_2\}$  in an action theory from whose dynamic part we (implicitly) infer  $\neg\varphi_2$ . In this case, contracting  $\neg\varphi_1$  while keeping  $\neg\varphi_2$  would necessarily ask for a change in  $\mathcal{S}$ .

We point out nevertheless that in both cases (modular and non-modular) the extra work in changing other modules stays in the mechanical level, i.e., in the algorithms that carry out the modification, and does not augment in a significant way the amount of work the knowledge engineer is expected to do. Moreover, considering the evolution of the theory, i.e., future modifications one should perform in it, modularity has to be checked/ensured only once, since it is preserved by our operators (cf. Lemma B.1).

While terminating, our algorithms come with a considerable computational cost: the entailment test in  $K_n$  with global axioms is known to be PSPACE-complete. Although this may be acceptable (theory change can be carried out offline), the computation of  $IP(\cdot)$  might result in exponential growth.

We have also extended Varzinczak's studies [52] by defining a semantics for action theory revision based on minimal modifications of models. For the corresponding revision algorithms, the reader is referred to the work by Varzinczak [53]. One of our ongoing researches is on assessing our revision operators' behavior w.r.t. the AGM postulates for revision [1].

Another issue that drives our future research on the subject is how to contract not only laws but *any* PDL-formula. As defined, the order of application of our operators matter in the final result: if we contract  $\varphi$  and then  $\varphi \rightarrow [a]\psi$  from a theory  $\mathcal{T}$ , the result may not be the same as contracting  $\varphi \rightarrow [a]\psi$  first and then removing  $\varphi$ . This problem would not appear in a more general framework in which any formula could be contracted: removing  $\varphi \wedge (\varphi \rightarrow [a]\psi)$  should give the same result as  $(\varphi \rightarrow [a]\psi) \wedge \varphi$ .

Definitions 3.1, 3.5 and 3.8 appear to be important for better understanding the problem of contracting general formulas: basically the set of modifications to perform in a given model in order to force it to falsify a general formula will comprise removal/addition of arrows/worlds. The definition of a general revision/contraction method will then benefit from our constructions.

Given the connection between multimodal logics and Description Logics [3], we believe that the definitions here given may also contribute to ontology evolution and debugging in DLs.

## Acknowledgements

The author is grateful to Arina Britz and Ken Halland for proofreading an earlier version of this article. Their comments helped very much in improving the text.

## References

- [1] C. Alchourrón, P. Gärdenfors, and D. Makinson. On the logic of theory change: Partial meet contraction and revision functions. *J. of Symbolic Logic*, 50:510–530, 1985.
- [2] E. Amir. (De)composition of situation calculus theories. In *Proc. 17th Natl. Conf. on Artificial Intelligence (AAAI’2000)*, pages 456–463, Austin, 2000. AAAI Press/MIT Press.
- [3] F. Baader, D. Calvanese, D. McGuinness, D. Nardi, and P. Patel-Schneider, editors. *Description Logic Handbook*. Cambridge University Press, 2003.
- [4] C. Baral and J. Lobo. Defeasible specifications in action theories. In M.E. Pollack, editor, *Proc. 15th Intl. Joint Conf. on Artificial Intelligence (IJCAI’97)*, pages 1441–1446, Nagoya, 1997. Morgan Kaufmann Publishers.
- [5] G. Brewka, S. Coradeschi, A. Perini, and P. Traverso, editors. *Proc. 17th Eur. Conf. on Artificial Intelligence (ECAI’06)*, Riva del Garda, 2006. IOS Press.
- [6] M. Castilho, O. Gasquet, and A. Herzig. Formalizing action and change in modal logic I: the frame problem. *J. of Logic and Computation*, 9(5):701–735, 1999.
- [7] L. Cholvy. Checking regulation consistency by using SOL-resolution. In *Proc. 7th Intl. Conf. on AI and Law*, pages 73–79, Oslo, 1999.

- [8] B. Cuenca Grau, B. Parsia, E. Sirin, and A. Kalyanpur. Modularity and web ontologies. In P. Doherty, J. Mylopoulos, and C. Welty, editors, *Proc. 10th Intl. Conf. on Knowledge Representation and Reasoning (KR'2006)*, pages 198–208, Lake District, 2006. Morgan Kaufmann Publishers.
- [9] M. Dalal. Investigations into a theory of knowledge base revision: preliminary report. In Smith and Mitchell [48], pages 475–479.
- [10] G. De Giacomo and M. Lenzerini. PDL-based framework for reasoning about actions. In M. Gori and G. Soda, editors, *Proc. 4th Congress of the Italian Association for Artificial Intelligence (IA\*AI'95)*, number 992 in LNAI, pages 103–114. Springer-Verlag, 1995.
- [11] R. Demolombe, A. Herzig, and I. Varzinczak. Regression in modal logic. *J. of Applied Non-Classical Logics (JANCL)*, 13(2):165–185, 2003.
- [12] T. Eiter, E. Erdem, M. Fink, and J. Senko. Updating action domain descriptions. In Kaelbling and Saffioti [29], pages 418–423.
- [13] T. Eiter, E. Erdem, M. Fink, and J. Senko. Resolving conflicts in action descriptions. In Brewka et al. [5], pages 367–371.
- [14] A. Fuhrmann. On the modal logic of theory change. In *The Logic of Theory Change*, pages 259–281, 1989.
- [15] P. Gärdenfors. *Knowledge in Flux: Modeling the Dynamics of Epistemic States*. MIT Press, Cambridge, MA, 1988.
- [16] M. Gelfond and V. Lifschitz. Representing action and change by logic programs. *Journal of Logic Programming*, 17(2/3&4):301–321, 1993.
- [17] E. Giunchiglia, G. Kartha, and V. Lifschitz. Representing action: indeterminacy and ramifications. *Artificial Intelligence*, 95(2):409–438, 1997.
- [18] R.W. Hamming. Error detecting and error correcting codes. *Bell System Technical Journal*, 26(2):147–160, 1950.
- [19] S. Hansson. *A Textbook of Belief Dynamics: Theory Change and Database Updating*. Kluwer Academic Publishers, 1999.
- [20] D. Harel, J. Tiuryn, and D. Kozen. *Dynamic Logic*. MIT Press, Cambridge, MA, 2000.

- [21] A. Herzig, L. Perrussel, and I. Varzinczak. Elaborating domain descriptions. In Brewka et al. [5], pages 397–401.
- [22] A. Herzig and O. Rifi. Propositional belief base update and minimal change. *Artificial Intelligence*, 115(1):107–138, 1999.
- [23] A. Herzig and I. Varzinczak. Domain descriptions should be modular. In R. López de Mántaras and L. Saitta, editors, *Proc. 16th Eur. Conf. on Artificial Intelligence (ECAI’04)*, pages 348–352, Valencia, 2004. IOS Press.
- [24] A. Herzig and I. Varzinczak. Cohesion, coupling and the meta-theory of actions. In Kaelbling and Saffiotti [29], pages 442–447.
- [25] A. Herzig and I. Varzinczak. On the modularity of theories. In R. Schmidt, I. Pratt-Hartmann, M. Reynolds, and H. Wansing, editors, *Advances in Modal Logic*, volume 5, pages 93–109. King’s College Publications, 2005. Selected papers of AiML 2004 (also available at <http://www.aiml.net/volumes/volume5>).
- [26] A. Herzig and I. Varzinczak. A modularity approach for a fragment of  $\mathcal{ALC}$ . In M. Fisher, W. van der Hoek, B. Konev, and A. Lisitsa, editors, *Proc. 10th Eur. Conf. on Logics in Artificial Intelligence (JELIA’2006)*, number 4160 in LNAI, pages 216–228. Springer-Verlag, 2006.
- [27] A. Herzig and I. Varzinczak. Metatheory of actions: beyond consistency. *Artificial Intelligence*, 171:951–984, 2007.
- [28] Y. Jin and M. Thielscher. Iterated belief revision, revised. In Kaelbling and Saffiotti [29], pages 478–483.
- [29] L. Kaelbling and A. Saffiotti, editors. *Proc. 19th Intl. Joint Conf. on Artificial Intelligence (IJCAI’05)*, Edinburgh, 2005. Morgan Kaufmann Publishers.
- [30] N. Kartha and V. Lifschitz. Actions with indirect effects (preliminary report). In J. Doyle, E. Sandewall, and P. Torasso, editors, *Proc. 4th Intl. Conf. on Knowledge Representation and Reasoning (KR’94)*, pages 341–350, Bonn, 1994. Morgan Kaufmann Publishers.
- [31] H. Katsuno and A. Mendelzon. On the difference between updating a knowledge base and revising it. In P. Gärdenfors, editor, *Belief revision*, pages 183–203. Cambridge University Press, 1992.

- [32] M. Kracht and F. Wolter. Properties of independently axiomatizable bimodal logics. *J. of Symbolic Logic*, 56(4):1469–1485, 1991.
- [33] R. Li and L.M. Pereira. What is believed is what is explained. In H. Shrobe and T. Senator, editors, *Proc. 13th Natl. Conf. on Artificial Intelligence (AAAI'96)*, pages 550–555, Portland, 1996. AAAI Press/MIT Press.
- [34] P. Liberatore. A framework for belief update. In *Proc. 7th Eur. Conf. on Logics in Artificial Intelligence (JELIA'2000)*, pages 361–375, 2000.
- [35] D. Makinson. Friendliness and sympathy in logic. In J.-Y. Beziau, editor, *Logica Universalis*. Springer-Verlag, 2nd edition, 2007.
- [36] P. Marquis. Consequence finding algorithms. In D. Gabbay and Ph. Smets, editors, *Algorithms for Defeasible and Uncertain Reasoning, in S. Moral, J. Kohlas (Eds), Handbook of Defeasible Reasoning and Uncertainty Management Systems*, volume 5, chapter 2, pages 41–145. Kluwer Academic Publishers, 2000.
- [37] J. McCarthy. Epistemological problems of artificial intelligence. In N. Sridharan, editor, *Proc. 5th Intl. Joint Conf. on Artificial Intelligence (IJCAI'77)*, pages 1038–1044, Cambridge, MA, 1977. Morgan Kaufmann Publishers.
- [38] J. McCarthy. Elaboration tolerance. In *Proc. Common Sense'98*, London, 1998. Available at <http://www-formal.stanford.edu/jmc/elaboration.html>.
- [39] J. McCarthy and P. Hayes. Some philosophical problems from the standpoint of artificial intelligence. In B. Meltzer and D. Mitchie, editors, *Machine Intelligence*, volume 4, pages 463–502. Edinburgh University Press, 1969.
- [40] B. Nebel. A knowledge level analysis of belief revision. In R. Brachman, H. Levesque, and R. Reiter, editors, *Proc. Intl. Conf. on Knowledge Representation and Reasoning (KR'89)*, pages 301–311, Toronto, 1989. Morgan Kaufmann Publishers.
- [41] R. Parikh. Beliefs, belief revision, and splitting languages. In L. Moss, editor, *Logic, Language and Computation*, volume 2 of *CSLI Lecture Notes*, pages 266–278. CSLI Publications, 1999.

- [42] S. Popkorn. *First Steps in Modal Logic*. Cambridge University Press, 1994.
- [43] W. V. O. Quine. The problem of simplifying truth functions. *American Mathematical Monthly*, 59:521–531, 1952.
- [44] W. V. O. Quine. Paradox. *Scientific American*, pages 84–96, 1962.
- [45] R. Reiter. The frame problem in the situation calculus: A simple solution (sometimes) and a completeness result for goal regression. In V. Lifschitz, editor, *Artificial Intelligence and Mathematical Theory of Computation: Papers in Honor of John McCarthy*, pages 359–380. Academic Press, San Diego, 1991.
- [46] S. Shapiro, Y. Lespérance, and H. Levesque. Goal change. In Kaelbling and Saffioti [29], pages 582–588.
- [47] S. Shapiro, M. Pagnucco, Y. Lespérance, and H. Levesque. Iterated belief change in the situation calculus. In T. Cohn, F. Giunchiglia, and B. Selman, editors, *Proc. 7th Intl. Conf. on Knowledge Representation and Reasoning (KR'2000)*, pages 527–538, Breckenridge, 2000. Morgan Kaufmann Publishers.
- [48] R. Smith and T. Mitchell, editors. *Proc. 7th Natl. Conf. on Artificial Intelligence (AAAI'88)*, St. Paul, 1988. Morgan Kaufmann Publishers.
- [49] M. Thielscher. Computing ramifications by postprocessing. In C. Mellish, editor, *Proc. 14th Intl. Joint Conf. on Artificial Intelligence (IJCAI'95)*, pages 1994–2000, Montreal, 1995. Morgan Kaufmann Publishers.
- [50] M. Thielscher. Ramification and causality. *Artificial Intelligence*, 89(1–2):317–364, 1997.
- [51] I. Varzinczak. *What is a good domain description? Evaluating and revising action theories in dynamic logic*. PhD thesis, Université Paul Sabatier, Toulouse, 2006.
- [52] I. Varzinczak. Action theory contraction and minimal change. In G. Brewka and J. Lang, editors, *Proc. 11th Intl. Conf. on Knowledge Representation and Reasoning (KR'2008)*, pages 651–661, Sydney, 2008. AAAI Press/MIT Press.

- [53] I. Varzinczak. Action theory revision in dynamic logic. In *Workshop on Nonmonotonic Reasoning (NMR'08)*, Sydney, 2008.
- [54] F. Veltman. Making counterfactual assumptions. *Journal of Semantics*, 22(2):159–180, 2005.
- [55] M.-A. Winslett. Reasoning about action using a possible models approach. In Smith and Mitchell [48], pages 89–93.
- [56] D. Zhang, S. Chopra, and N. Foo. Consistency of action descriptions. In M. Ishizuka and A. Sattar, editors, *Proc. 7th Pacific Rim Intl. Conf. on Artificial Intelligence: Trends in Artificial Intelligence*, number 2417 in LNCS, pages 70–79. Springer-Verlag, 2002.
- [57] D. Zhang and N. Foo. EPDL: A logic for causal reasoning. In B. Nebel, editor, *Proc. 17th Intl. Joint Conf. on Artificial Intelligence (IJCAI'01)*, pages 131–138, Seattle, 2001. Morgan Kaufmann Publishers.

## A Proof of Theorem 5.2

Let  $\mathcal{T}$  be modular, and  $\Phi$  be a law. For all  $\mathcal{M}' \in \mathcal{M}_{\Phi}^-$  such that  $\models^{\mathcal{M}} \mathcal{T}$  for every  $\mathcal{M} \in \mathcal{M}$ , there is  $\mathcal{T}' \in \mathcal{T}_{\Phi}^-$  such that  $\models^{\mathcal{M}'} \mathcal{T}'$  for every  $\mathcal{M}' \in \mathcal{M}'$ .

**Lemma A.1**  $\mathcal{T} \models_{\text{PDL}} \mathcal{T}'$ .

**Proof:** Let  $\mathcal{T}$  be an action theory, and  $\mathcal{T}' \in \mathcal{T}_{\Phi}^-$ , for  $\Phi$  a law. We analyze each case.

Let  $\Phi$  be of the form  $\varphi \rightarrow \langle a \rangle \top$ , for some  $\varphi \in \mathfrak{Fml}$ . Then  $\mathcal{T}'$  is such that

$$\mathcal{T}' = (\mathcal{T} \setminus \mathcal{X}_a) \cup \{(\varphi_i \wedge \neg(\pi \wedge \varphi_A)) \rightarrow \langle a \rangle \top : \varphi_i \rightarrow \langle a \rangle \top \in \mathcal{X}_a\}$$

where  $\pi \in IP(\mathcal{S} \wedge \varphi)$  and  $\varphi_A = \bigwedge_{\substack{p_i \in atm(\pi) \\ p_i \in A}} p_i \wedge \bigwedge_{\substack{p_i \in atm(\pi) \\ p_i \notin A}} \neg p_i$ , for some  $A \subseteq \overline{atm(\pi)}$ .

Let  $\mathcal{M} = \langle W, R \rangle$  be such that  $\models^{\mathcal{M}} \mathcal{T}$ . It is enough to show that  $\mathcal{M}$  is a model of the new laws. For every  $(\varphi_i \wedge \neg(\pi \wedge \varphi_A)) \rightarrow \langle a \rangle \top$ , for every  $w \in W$ , if  $\models_w^{\mathcal{M}} \varphi_i \wedge \neg(\pi \wedge \varphi_A)$ , then  $\models_w^{\mathcal{M}} \varphi_i$ . Because  $\mathcal{T} \models_{\text{PDL}} \varphi_i \rightarrow \langle a \rangle \top$ ,  $\models^{\mathcal{M}} \varphi_i \rightarrow \langle a \rangle \top$ , and then  $R_a(w) \neq \emptyset$ .

Hence  $\models^{\mathcal{M}} \mathcal{T}'$ .



Let now  $\Phi$  have the form  $\varphi \rightarrow [a]\psi$ , for  $\varphi, \psi \in \mathfrak{Fml}$ . Then  $\mathcal{T}'$  is such that

$$\begin{aligned} & (\mathcal{T} \setminus \mathcal{E}_a^-) \cup \\ & \{(\varphi_i \wedge \neg(\pi \wedge \varphi_A)) \rightarrow [a]\psi_i : \varphi_i \rightarrow [a]\psi_i \in \mathcal{E}_a^-\} \cup \\ \mathcal{T}' = & \{(\varphi_i \wedge \pi \wedge \varphi_A) \rightarrow [a](\psi_i \vee \pi') : \varphi_i \rightarrow [a]\psi_i \in \mathcal{E}_a^-\} \cup \\ & \left\{ \begin{array}{l} \ell \in L, \text{ for some } L \subseteq \mathfrak{Lit} \text{ s.t.} \\ (\pi \wedge \varphi_A \wedge \ell) \rightarrow [a](\psi \vee \ell) : \mathcal{S} \not\models (\pi' \wedge \bigwedge_{\ell \in L} \ell) \rightarrow \perp, \text{ and } \ell \in \pi' \\ \text{or } \mathcal{T} \not\models_{\text{PDL}} (\pi \wedge \varphi_A \wedge \ell) \rightarrow [a]\neg\ell \end{array} \right\} \end{aligned}$$

where  $\mathcal{E}_a^- = \bigcup_{1 \leq i \leq n} (\mathcal{E}_a^{\varphi, \psi})_i$ ,  $\pi \in IP(\mathcal{S} \wedge \varphi)$ ,  $\varphi_A = \bigwedge_{\substack{p_i \in \overline{atm(\pi)}}} p_i \wedge \bigwedge_{\substack{p_i \in \overline{atm(\pi)}}} \neg p_i$ ,  
for some  $A \subseteq \overline{atm(\pi)}$ , and  $\pi' \in IP(\mathcal{S} \wedge \neg\psi)$ .

Let  $\mathcal{M} = \langle W, R \rangle$  be such that  $\models^{\mathcal{M}} \mathcal{T}$ . It is enough to show that  $\mathcal{M}$  is a model of the added laws. Given  $(\varphi_i \wedge \neg(\pi \wedge \varphi_A)) \rightarrow [a]\psi_i$ , for every  $w \in W$ , if  $\models_w^{\mathcal{M}} \varphi_i \wedge \neg(\pi \wedge \varphi_A)$ , then  $\models_w^{\mathcal{M}} \varphi_i$ . Because  $\mathcal{T} \models_{\text{PDL}} \varphi_i \rightarrow [a]\psi_i$ ,  $\models^{\mathcal{M}} \varphi_i \rightarrow [a]\psi_i$ , and then  $\models_{w'}^{\mathcal{M}} \psi_i$  for every  $w' \in W$  such that  $(w, w') \in R_a$ .

For  $(\varphi_i \wedge \pi \wedge \varphi_A) \rightarrow [a](\psi_i \vee \pi')$ , for every  $w \in W$ , if  $\models_w^{\mathcal{M}} \varphi_i \wedge \pi \wedge \varphi_A$ , then again  $\models_{w'}^{\mathcal{M}} \psi_i$  for every  $w' \in W$  such that  $(w, w') \in R_a$ .

Now, given  $(\pi \wedge \varphi_A \wedge \ell) \rightarrow [a](\psi \vee \ell)$ , for every  $w \in W$ , if  $\models_w^{\mathcal{M}} \pi \wedge \varphi_A \wedge \ell$ , then  $\models_w^{\mathcal{M}} \pi$ , and then  $\models_w^{\mathcal{M}} \varphi$ . Since  $\mathcal{T} \models_{\text{PDL}} \varphi \rightarrow [a]\psi$ , we have  $\models^{\mathcal{M}} \varphi \rightarrow [a]\psi$ , and then  $\models_{w'}^{\mathcal{M}} \psi$  for every  $w' \in W$  such that  $(w, w') \in R_a$ .

Hence  $\models^{\mathcal{M}} \mathcal{T}'$ .

Let  $\Phi$  be a propositional  $\varphi$ . Then  $\mathcal{T}'$  is such that

$$\begin{aligned} & ((\mathcal{T} \setminus \mathcal{S}) \cup \mathcal{S}^-) \setminus \mathcal{X}_a \cup \\ \mathcal{T}' = & \{(\varphi_i \wedge \varphi) \rightarrow \langle a \rangle \top : \varphi_i \rightarrow \langle a \rangle \top \in \mathcal{X}_a\} \cup \\ & \{\neg\varphi \rightarrow [a]\perp\} \end{aligned}$$

for some  $\mathcal{S}^- \in \mathcal{S} \ominus \varphi$ .

Let  $\mathcal{M} = \langle W, R \rangle$  be such that  $\models^{\mathcal{M}} \mathcal{T}$ . It suffices to show that  $\mathcal{M}$  satisfies the added laws.

Since we assume  $\ominus$  behaves like a classical contraction operator, like e.g. Katsuno and Mendelzon's [31], we have  $\models_{\text{CPL}} \mathcal{S} \rightarrow \mathcal{S}^-$ , and then, because  $\models^{\mathcal{M}} \mathcal{S}$ , we have  $\models^{\mathcal{M}} \mathcal{S}^-$ .

Now given  $(\varphi_i \wedge \varphi) \rightarrow \langle a \rangle \top$ , for every  $w \in W$ , if  $\models_w^{\mathcal{M}} \varphi_i \wedge \varphi$ , then  $\models_w^{\mathcal{M}} \varphi_i$ , and because  $\models \varphi_i \rightarrow \langle a \rangle \top$ , we have  $R_a(w) \neq \emptyset$ .  
 Finally, for  $\neg\varphi \rightarrow [a]\perp$ , because  $\models \varphi$ ,  $\mathcal{M}$  trivially satisfies  $\neg\varphi \rightarrow [a]\perp$ .  
 Hence,  $\models^{\mathcal{M}} T'$ .  $\blacksquare$

## Proof of Theorem 5.2

Let  $\mathcal{M} = \{\mathcal{M} : \models^{\mathcal{M}} T\}$ , and  $\mathcal{M}' \in \mathcal{M}_{\Phi}^-$ . We show that there is  $T' \in \mathcal{T}_{\Phi}^-$  such that  $\models^{\mathcal{M}'} T'$  for every  $\mathcal{M}' \in \mathcal{M}'$ .

By definition, each  $\mathcal{M}' \in \mathcal{M}'$  is such that either  $\models^{\mathcal{M}'} T$  or  $\not\models^{\mathcal{M}'} \Phi$ . Because  $\mathcal{T}_{\Phi}^- \neq \emptyset$ , there must be  $T' \in \mathcal{T}_{\Phi}^-$ . If  $\models^{\mathcal{M}'} T$ , by Lemma A.1  $\models^{\mathcal{M}'} T'$  and we are done. Let's then suppose that  $\not\models^{\mathcal{M}'} \Phi$ . We analyze each case.

Let  $\Phi$  have the form  $\varphi \rightarrow \langle a \rangle \top$  for some  $\varphi \in \mathfrak{Fml}$ . Then  $\mathcal{M}' = \langle W', R' \rangle$ , where  $W' = W$ ,  $R' = R \setminus R_a^{\varphi}$ , with  $R_a^{\varphi} = \{(w, w') : \models_w^{\mathcal{M}} \varphi \text{ and } (w, w') \in R_a\}$ , for some  $\mathcal{M} \in \mathcal{M}$ .

Let  $u \in W'$  be such that  $\not\models_u^{\mathcal{M}'} \varphi \rightarrow \langle a \rangle \top$ , i.e.,  $\models_u^{\mathcal{M}'} \varphi$  and  $R'_a(u) = \emptyset$ .

Because  $u \Vdash \varphi$ , there must be  $v \in \text{base}(\varphi, W')$  such that  $v \subseteq u$ . Let  $\pi = \bigwedge_{\ell \in v} \ell$ . Clearly  $\pi$  is a prime implicant of  $\mathcal{S} \wedge \varphi$ . Let also  $\varphi_A = \bigwedge_{\ell \in u \setminus v} \ell$ , and consider

$$T' = (T \setminus \mathcal{X}_a) \cup \{(\varphi_i \wedge \neg(\pi \wedge \varphi_A)) \rightarrow \langle a \rangle \top : \varphi_i \rightarrow \langle a \rangle \top \in \mathcal{X}_a\}$$

(Clearly,  $T'$  is a theory produced by Algorithm 1.)

It is enough to show that  $\mathcal{M}'$  is a model of the new added laws. Given  $(\varphi_i \wedge \neg(\pi \wedge \varphi_A)) \rightarrow \langle a \rangle \top \in T'$ , for every  $w \in W'$ , if  $\models_w^{\mathcal{M}'} \varphi_i \wedge \neg(\pi \wedge \varphi_A)$ , then  $\models_w^{\mathcal{M}'} \varphi_i$ , from what it follows  $\models_w^{\mathcal{M}} \varphi_i$ . Because  $\models \varphi_i \rightarrow \langle a \rangle \top$ , there is  $w' \in W$  such that  $w' \in R_a(w)$ . We need to show that  $(w, w') \in R'_a$ . If  $\not\models_w^{\mathcal{M}} \varphi$ , then  $R_a^{\varphi} = \emptyset$ , and  $(w, w') \in R'_a$ . If  $\models_w^{\mathcal{M}} \varphi$ , either  $w = u$ , and then from  $\models_u^{\mathcal{M}'} \pi \wedge \varphi_A$  we conclude  $\models_u^{\mathcal{M}'} (\varphi_i \wedge \neg(\pi \wedge \varphi_A)) \rightarrow \langle a \rangle \top$ , or  $w \neq u$  and then we must have  $(w, w') \in R'_a$ , otherwise there is  $S_a^{\varphi} \subset R_a^{\varphi}$  such that  $R \setminus (R \setminus S_a^{\varphi}) \subset R \setminus (R \setminus R_a^{\varphi})$ , and then  $\mathcal{M}'' = \langle W', R \setminus S_a^{\varphi} \rangle$  is such that  $\not\models^{\mathcal{M}''} \varphi \rightarrow \langle a \rangle \top$  and  $\mathcal{M}'' \preceq_{\mathcal{M}} \mathcal{M}'$ , a contradiction because  $\mathcal{M}'$  is minimal w.r.t.  $\preceq_{\mathcal{M}}$ . Thus  $(w, w') \in R'_a$ , and then  $\models_w^{\mathcal{M}'} \langle a \rangle \top$ . Hence  $\models^{\mathcal{M}'} T'$ .

Now let  $\Phi$  be of the form  $\varphi \rightarrow [a]\psi$ , for  $\varphi, \psi$  both Boolean. Then  $\mathcal{M}' = \langle W', R' \rangle$ , where  $W' = W$ ,  $R' = R \cup R_a^{\varphi, \neg\psi}$ , with

$$R_a^{\varphi, \neg\psi} = \{(w, w') : w' \in \text{RelTarget}(w, \varphi \rightarrow [a]\psi, \mathcal{M}, \mathcal{M})\}$$

for some  $\mathcal{M} = \langle W, R \rangle \in \mathcal{M}$ .

Let  $u \in W'$  be such that  $\not\models_u^{\mathcal{M}'} \varphi \rightarrow [a]\psi$ . Then there is  $u' \in W'$  such that  $(u, u') \in R'_a$  and  $\not\models_{u'}^{\mathcal{M}'} \psi$ . Because  $u \models \varphi$ , there is  $v \in \text{base}(\varphi, W')$  such that  $v \subseteq u$ , and as  $u' \models \neg\psi$ , there must be  $v' \in \text{base}(\neg\psi, W')$  such that  $v' \subseteq u'$ . Let  $\pi = \bigwedge_{\ell \in v} \ell$ ,  $\varphi_A = \bigwedge_{\ell \in u} \ell$ , and  $\pi' = \bigwedge_{\ell \in v'} \ell$ . Clearly  $\pi$  (resp.  $\pi'$ ) is a prime implicant of  $\mathcal{S} \wedge \varphi$  (resp.  $\mathcal{S} \wedge \neg\psi$ ).

Now let  $\mathcal{E}_a^- = \bigcup_{1 \leq i \leq n} (\mathcal{E}_a^{\varphi, \psi})_i$  and let the theory

$$\begin{aligned} & (T \setminus \mathcal{E}_a^-) \cup \\ & \{(\varphi_i \wedge \neg(\pi \wedge \varphi_A)) \rightarrow [a]\psi_i : \varphi_i \rightarrow [a]\psi_i \in \mathcal{E}_a^-\} \cup \\ \mathcal{T}' = & \{(\varphi_i \wedge \pi \wedge \varphi_A) \rightarrow [a](\psi_i \vee \pi') : \varphi_i \rightarrow [a]\psi_i \in \mathcal{E}_a^-\} \cup \\ & \left\{ \begin{array}{l} \ell \in L, \text{ for some } L \subseteq \mathfrak{Lit} \text{ s.t.} \\ (\pi \wedge \varphi_A \wedge \ell) \rightarrow [a](\psi \vee \ell) : \mathcal{S} \not\models (\pi' \wedge \bigwedge_{\ell \in L} \ell) \rightarrow \perp, \text{ and } \ell \in \pi' \\ \text{or } T \not\models_{\text{PDL}} (\pi \wedge \varphi_A \wedge \ell) \rightarrow [a]\neg\ell \end{array} \right\} \end{aligned}$$

(Clearly,  $\mathcal{T}'$  is a theory produced by Algorithm 2.)

In order to show that  $\mathcal{M}'$  is a model of  $\mathcal{T}'$ , it is enough to show that it is a model of the added laws. Given  $(\varphi_i \wedge \neg(\pi \wedge \varphi_A)) \rightarrow [a]\psi_i \in \mathcal{T}'$ , for every  $w \in W'$ , if  $\models_w^{\mathcal{M}'} \varphi_i \wedge \neg(\pi \wedge \varphi_A)$ , then  $\models_w^{\mathcal{M}'} \varphi_i$ , and then  $\models_w^{\mathcal{M}} \varphi_i$ . Because  $\models^{\mathcal{M}} \varphi_i \rightarrow [a]\psi_i$ ,  $\models_{w'}^{\mathcal{M}} \psi_i$  for all  $w' \in W$  such that  $(w, w') \in R_a$ . We need to show that  $R'_a(w) = R_a(w)$ . If  $\not\models_w^{\mathcal{M}'} \varphi$ , then  $R_a^{\varphi, \neg\psi} = \emptyset$ , and then  $R'_a(w) = R_a(w)$ . If  $\models_w^{\mathcal{M}'} \varphi$ , then either  $w = u$ , and from  $\models_u^{\mathcal{M}'} \pi \wedge \varphi_A$  we conclude  $\models_u^{\mathcal{M}'} (\varphi_i \wedge \neg(\pi \wedge \varphi_A)) \rightarrow [a]\psi_i$ , or  $w \neq u$ , and then we must have  $R_a^{\varphi, \neg\psi} = \emptyset$ , otherwise there would be  $S_a^{\varphi, \neg\psi} \subset R_a^{\varphi, \neg\psi}$  such that  $R^-(R \cup S_a^{\varphi, \neg\psi}) \subset R^-(R \cup R_a^{\varphi, \neg\psi})$ , and then  $\mathcal{M}'' = \langle W', R \cup S_a^{\varphi, \neg\psi} \rangle$  would be such that  $\not\models^{\mathcal{M}''} \varphi \rightarrow [a]\psi$  and  $\mathcal{M}'' \preceq_{\mathcal{M}} \mathcal{M}'$ , a contradiction since  $\mathcal{M}'$  is minimal w.r.t.  $\preceq_{\mathcal{M}}$ . Hence  $R'_a(w) = R_a(w)$ , and  $\models_{w'}^{\mathcal{M}'} \psi_i$  for all  $w'$  such that  $(w, w') \in R'_a$ .

Now, given  $(\varphi_i \wedge \pi \wedge \varphi_A) \rightarrow [a](\psi_i \vee \pi')$ , for every  $w \in W'$ , if  $\models_w^{\mathcal{M}'} \varphi_i \wedge \pi \wedge \varphi_A$ , then  $\models_w^{\mathcal{M}'} \varphi_i$ , and then  $\models_w^{\mathcal{M}} \varphi_i$ . Because,  $\models^{\mathcal{M}} \varphi_i \rightarrow [a]\psi_i$ , we have  $\models_{w'}^{\mathcal{M}} \psi_i$  for all

$w' \in W$  such that  $(w, w') \in R_a$ , and then  $\models_{w'}^{\mathcal{M}'} \psi_i$  for every  $w' \in W'$  such that  $(w, w') \in R'_a \setminus R_a^{\varphi, \neg\psi}$ . Now, given  $(w, w') \in R_a^{\varphi, \neg\psi}$ ,  $\models_{w'}^{\mathcal{M}'} \pi'$ , and the result follows.

Now, for each  $(\pi \wedge \varphi_A \wedge \ell) \rightarrow [a](\psi \vee \ell)$ , for every  $w \in W'$ , if  $\models_w^{\mathcal{M}'} \pi \wedge \varphi_A \wedge \ell$ , then  $\models_w^{\mathcal{M}'} \varphi$ , and then  $\models_w^{\mathcal{M}} \varphi$ . Because  $\models_w^{\mathcal{M}} \varphi \rightarrow [a]\psi$ , we have  $\models_w^{\mathcal{M}} \psi$  for every  $w' \in W$  such that  $(w, w') \in R_a$ , and then  $\models_{w'}^{\mathcal{M}'} \psi$  for all  $w' \in W'$  such that  $(w, w') \in R'_a \setminus R_a^{\varphi, \neg\psi}$ . It remains to show that  $\models_{w'}^{\mathcal{M}'} \ell$  for every  $w' \in W'$  such that  $(w, w') \in R_a^{\varphi, \neg\psi}$ . Since  $\mathcal{M}'$  is minimal, it is enough to show that  $\models_u^{\mathcal{M}'} \ell$  for every  $\ell \in \mathfrak{Lit}$  such that  $\models_u^{\mathcal{M}'} \pi \wedge \varphi_A \wedge \ell$ . If  $\ell \in \pi'$ , the result follows. Otherwise, suppose  $\not\models_{u'}^{\mathcal{M}'} \ell$ . Then

- either  $\neg\ell \in \pi'$ , then  $\pi'$  and  $\ell$  are unsatisfiable, and in this case Algorithm 2 has not put the law  $(\pi \wedge \varphi_A \wedge \ell) \rightarrow [a](\psi \vee \ell)$  in  $\mathcal{T}'$ , a contradiction;
- or  $\neg\ell \in u' \setminus v'$ . In this case, there is a valuation  $u'' = (u' \setminus \{\neg\ell\}) \cup \{\ell\}$  such that  $u'' \not\models \psi$ . We must have  $u'' \in W'$ , otherwise there will be  $L' = \{\ell_i : \ell_i \in u''\}$  such that  $\mathcal{T} \models_{\text{PDL}} (\pi' \wedge \bigwedge_{\ell_i \in L'} \ell_i) \rightarrow \perp$ , and, because  $\mathcal{T}$  is modular,  $\mathcal{S} \models_{\text{CPL}} (\pi' \wedge \bigwedge_{\ell_i \in L'} \ell_i) \rightarrow \perp$ , and then Algorithm 2 has not put the law  $(\pi \wedge \varphi_A \wedge \ell) \rightarrow [a](\psi \vee \ell)$  in  $\mathcal{T}'$ , a contradiction. Then  $u'' \in W'$ , and moreover  $u'' \notin R_a^{\varphi, \neg\psi}(u)$ , otherwise  $\mathcal{M}'$  is not minimal. As  $u'' \setminus u \subset u' \setminus u$ , the only reason why  $u'' \notin R_a^{\varphi, \neg\psi}(u)$  is that there is  $\ell' \in u \cap u''$  such that  $\models_{\mathcal{M}_i} \bigwedge_{\ell_j \in u} \ell_j \rightarrow [a]\neg\ell'$  for every  $\mathcal{M}_i \in \mathcal{M}$  if and only if  $\ell' \notin v'$  for any  $v' \in \text{base}(\neg\psi, W')$  such that  $v' \subseteq u''$ . Clearly  $\ell' = \ell$ , and because  $\ell \notin \pi'$ , we have  $\models_{\mathcal{M}_i} \bigwedge_{\ell_j \in u} \ell_j \rightarrow [a]\neg\ell$  for every  $\mathcal{M}_i \in \mathcal{M}$ . Then  $\mathcal{T} \models_{\text{PDL}} (\pi \wedge \varphi_A \wedge \ell) \rightarrow [a]\neg\ell$ , and then Algorithm 2 has not put the law  $(\pi \wedge \varphi_A \wedge \ell) \rightarrow [a](\psi \vee \ell)$  in  $\mathcal{T}'$ , a contradiction.

Hence we have  $\models_{w'}^{\mathcal{M}'} \psi \vee \ell$  for every  $w' \in W'$  such that  $(w, w') \in R'_a$ .

Putting the above results together, we get  $\models^{\mathcal{M}'} \mathcal{T}'$ .

Let now  $\Phi$  be some propositional  $\varphi$ . Then  $\mathcal{M}' = \langle W', R' \rangle$ , where  $W \subseteq W'$ ,  $R' = R$ , is minimal w.r.t.  $\preceq_{\mathcal{M}}$ , i.e.,  $W'$  is a minimum superset of  $W$  such that there is  $u \in W'$  with  $u \not\models \varphi$ . Because we have assumed the syntactical classical contraction operator is sound and complete w.r.t. its semantics and is moreover minimal, then there must be  $\mathcal{S}^- \in \mathcal{S} \ominus \varphi$  such that  $W' = \text{val}(\mathcal{S}^-)$ . Hence  $\models^{\mathcal{M}'} \mathcal{S}^-$ .

Because  $R' = R$ , every effect law of  $\mathcal{T}$  remains true in  $\mathcal{M}'$ .  
Now, let

$$\begin{aligned} \mathcal{T}' = & ((\mathcal{T} \setminus \mathcal{S}) \cup \mathcal{S}^-) \setminus \mathcal{X}_a \cup \\ & \{(\varphi_i \wedge \varphi) \rightarrow \langle a \rangle \top : \varphi_i \rightarrow \langle a \rangle \top \in \mathcal{X}_a\} \cup \\ & \{\neg \varphi \rightarrow [a] \perp\} \end{aligned}$$

(Clearly,  $\mathcal{T}'$  is a theory produced by Algorithm 3.)

For every  $(\varphi_i \wedge \varphi) \rightarrow \langle a \rangle \top \in \mathcal{T}'$  and every  $w \in W'$ , if  $\models_w^{\mathcal{M}'} \varphi_i \wedge \varphi$ , then  $R_a(w) \neq \emptyset$ , because  $\models_w^{\mathcal{M}} \varphi_i \rightarrow \langle a \rangle \top$ . Given  $\neg \varphi \rightarrow [a] \perp$ , for every  $w \in W'$ , if  $\models_w^{\mathcal{M}'} \neg \varphi$ , then  $w = u$ , and  $R_a(w) = \emptyset$ .

Putting all these results together, we have  $\models^{\mathcal{M}'} \mathcal{T}'$ . ■

## B Proof of Theorem 5.3

Let  $\mathcal{T}$  be modular,  $\Phi$  a law, and  $\mathcal{T}' \in \mathcal{T}_{\Phi}^-$ . For all  $\mathcal{M}'$  such that  $\models^{\mathcal{M}'} \mathcal{T}'$ , there is  $\mathcal{M}' \in \mathcal{M}_{\Phi}^-$  such that  $\mathcal{M}' \in \mathcal{M}'$  and  $\models^{\mathcal{M}} \mathcal{T}$  for every  $\mathcal{M} \in \mathcal{M}$ .

**Lemma B.1** *Let  $\Phi$  be a law. If  $\mathcal{T}$  is modular, then every  $\mathcal{T}' \in \mathcal{T}_{\Phi}^-$  is modular.*

**Proof:** Let  $\Phi$  be nonclassical, and suppose there is  $\mathcal{T}' \in \mathcal{T}_{\Phi}^-$  such that  $\mathcal{T}'$  is not modular. Then there is some  $\varphi' \in \mathfrak{Fml}$  such that  $\mathcal{T}' \models_{\text{PDL}} \varphi'$  and  $\mathcal{S}' \not\models_{\text{CPL}} \varphi'$ , where  $\mathcal{S}'$  is the set of static laws in  $\mathcal{T}'$ . By Lemma A.1,  $\mathcal{T} \models_{\text{PDL}} \mathcal{T}'$ , and then we have  $\mathcal{T} \models_{\text{PDL}} \varphi'$ . Because  $\Phi$  is nonclassical,  $\mathcal{S}' = \mathcal{S}$ . Thus  $\mathcal{S} \not\models_{\text{CPL}} \varphi'$ , and hence  $\mathcal{T}$  is not modular.

Let now  $\Phi$  be some  $\varphi \in \mathfrak{Fml}$ . Then

$$\begin{aligned} \mathcal{T}' = & ((\mathcal{T} \setminus \mathcal{S}) \cup \mathcal{S}^-) \setminus \mathcal{X}_a \cup \\ & \{(\varphi_i \wedge \varphi) \rightarrow \langle a \rangle \top : \varphi_i \rightarrow \langle a \rangle \top \in \mathcal{X}_a\} \cup \\ & \{\neg \varphi \rightarrow [a] \perp\} \end{aligned}$$

for some  $\mathcal{S}^- \in \mathcal{S} \ominus \varphi$ .

Suppose  $\mathcal{T}$  is modular, and let  $\varphi' \in \mathfrak{Fml}$  be such that  $\mathcal{T}' \models_{\text{PDL}} \varphi'$  and  $\mathcal{S}^- \not\models_{\text{CPL}} \varphi'$ .

As  $\mathcal{S}^- \not\models_{\text{CPL}} \varphi'$ , there is  $v \in \text{val}(\mathcal{S}^-)$  such that  $v \not\models \varphi'$ . If  $v \in \text{val}(\mathcal{S})$ , then  $\mathcal{S} \not\models_{\text{CPL}} \varphi'$ , and as  $\mathcal{T}$  is modular,  $\mathcal{T} \not\models_{\text{PDL}} \varphi'$ . By Lemma A.1,  $\mathcal{T} \models_{\text{PDL}} \mathcal{T}'$ , and we

have  $\mathcal{T}' \not\models_{\text{PDL}} \varphi'$ , a contradiction. Hence  $v \notin \text{val}(\mathcal{S})$ . Moreover, we must have  $v \not\models \varphi$ , otherwise  $\ominus$  has not worked as expected.

Let  $\mathcal{M} = \langle W, R \rangle$  be such that  $\models^{\mathcal{M}} \mathcal{T}'$ . (We extend  $\mathcal{M}$  to another model of  $\mathcal{T}'$ .) Let  $\mathcal{M}' = \langle W', R' \rangle$  be such that  $W' = W \cup \{v\}$  and  $R' = R$ . To show that  $\mathcal{M}'$  is a model of  $\mathcal{T}'$ , it suffices to show that  $v$  satisfies every law in  $\mathcal{T}'$ . As  $v \in \text{val}(\mathcal{S}^-)$ ,  $\models_v^{\mathcal{M}'} \mathcal{S}^-$ . Given  $\neg\varphi \rightarrow [a]\perp \in \mathcal{T}'$ , as  $v \not\models \varphi$  and  $R'_a(v) = \emptyset$ ,  $\models_v^{\mathcal{M}'} \neg\varphi \rightarrow [a]\perp$ . Now, for every  $\varphi_i \rightarrow [a]\psi_i \in \mathcal{T}'$ , if  $\models_v^{\mathcal{M}'} \varphi_i$ , then we trivially have  $\models_v^{\mathcal{M}'} \psi_i$  for every  $v'$  such that  $(v, v') \in R'_a$ . Finally, given  $(\varphi_i \wedge \varphi) \rightarrow \langle a \rangle \top \in \mathcal{T}'$ , as  $v \not\models \varphi$ , the formula trivially holds in  $v$ . Hence  $\models^{\mathcal{M}'} \mathcal{T}'$ , and because there is  $v \in W'$  such that  $\not\models_v^{\mathcal{M}'} \varphi'$ , we have  $\mathcal{T}' \not\models_{\text{PDL}} \varphi'$ , a contradiction. Hence for all  $\varphi' \in \mathfrak{Fml}$  such that  $\mathcal{T}' \models_{\text{PDL}} \varphi'$ ,  $\mathcal{S}^- \models_{\text{CPL}} \varphi'$ , and then  $\mathcal{T}'$  is modular. ■

**Lemma B.2** *If  $\mathcal{M}_{big} = \langle W_{big}, R_{big} \rangle$  is a model of  $\mathcal{T}$ , then for every  $\mathcal{M} = \langle W, R \rangle$  such that  $\models^{\mathcal{M}} \mathcal{T}$  there is a minimal (w.r.t. set inclusion) extension  $R' \subseteq R_{big} \setminus R$  such that  $\mathcal{M}' = \langle \text{val}(\mathcal{S}), R \cup R' \rangle$  is a model of  $\mathcal{T}$ .*

**Proof:** Let  $\mathcal{M}_{big} = \langle W_{big}, R_{big} \rangle$  be a model of  $\mathcal{T}$ , and let  $\mathcal{M} = \langle W, R \rangle$  be such that  $\models^{\mathcal{M}} \mathcal{T}$ . Consider  $\mathcal{M}' = \langle \text{val}(\mathcal{S}), R \rangle$ . If  $\models^{\mathcal{M}'} \mathcal{T}$ , we have  $R' = \emptyset \subseteq R_{big} \setminus R$  that is minimal. Suppose then  $\not\models^{\mathcal{M}'} \mathcal{T}$ . We extend  $\mathcal{M}'$  to a model of  $\mathcal{T}$  that is a minimal extension of  $\mathcal{M}$ . As  $\not\models^{\mathcal{M}'} \mathcal{T}$ , there is  $v \in \text{val}(\mathcal{S}) \setminus W$  such that  $\not\models_v^{\mathcal{M}'} \mathcal{T}$ . Then there is  $\Phi \in \mathcal{T}$  such that  $\not\models_v^{\mathcal{M}'} \Phi$ . If  $\Phi$  is some  $\varphi \in \mathfrak{Fml}$ , as  $v \in W_{big}$ ,  $\mathcal{M}_{big}$  is not a model of  $\mathcal{T}$ . If  $\Phi$  is of the form  $\varphi \rightarrow [a]\psi$ , for  $\varphi, \psi \in \mathfrak{Fml}$ , there is  $v' \in \text{val}(\mathcal{S})$  such that  $(v, v') \in R_a$  and  $v' \not\models \psi$ , a contradiction since  $R_a(v) = \emptyset$ . Let now  $\Phi$  have the form  $\varphi \rightarrow \langle a \rangle \top$  for some  $\varphi \in \mathfrak{Fml}$ . Then  $\models_v^{\mathcal{M}'} \varphi$ . As  $v \in W_{big}$ , if  $\models_v^{\mathcal{M}_{big}} \varphi \rightarrow \langle a \rangle \top$ , then  $\models_v^{\mathcal{M}_{big}} \mathcal{T}$ . Hence,  $R_{big_a}(v) \neq \emptyset$ . Thus taking any  $(v, v') \in R_{big_a}$  gives us a minimal  $R' = \{(v, v')\}$  such that  $\mathcal{M}'' = \langle \text{val}(\mathcal{S}), R \cup R' \rangle$  is a model of  $\mathcal{T}$ . ■

**Lemma B.3** *Let  $\mathcal{T}$  be modular, and  $\Phi$  be a law. Then  $\mathcal{T} \models_{\text{PDL}} \Phi$  if and only if every  $\mathcal{M}' = \langle \text{val}(\mathcal{S}), R' \rangle$  such that  $\models^{(W, R)} \mathcal{T}$  and  $R \subseteq R'$  is a model of  $\Phi$ .*

**Proof:**

( $\Rightarrow$ ): Straightforward, as  $\mathcal{T} \models_{\text{PDL}} \Phi$  implies  $\models^{\mathcal{M}} \Phi$  for every  $\mathcal{M}$  such that  $\models^{\mathcal{M}} \mathcal{T}$ , in particular for those that are extensions of some model of  $\mathcal{T}$ .

( $\Leftarrow$ ): Suppose  $\mathcal{T} \not\models_{\text{PDL}} \Phi$ . Then there is  $\mathcal{M} = \langle W, R \rangle$  such that  $\models^{\mathcal{M}} \mathcal{T}$  and  $\not\models^{\mathcal{M}} \Phi$ . As  $\mathcal{T}$  is modular, the big model  $\mathcal{M}_{big} = \langle W_{big}, R_{big} \rangle$  of  $\mathcal{T}$  is a model of  $\mathcal{T}$ . Then by Lemma B.2 there is a minimal extension  $R'$  of  $R$  w.r.t.  $R_{big}$  such that  $\mathcal{M}' = \langle \text{val}(\mathcal{S}), R \cup R' \rangle$  is a model of  $\mathcal{T}$ . Because  $\not\models^{\mathcal{M}} \Phi$ , there is  $w \in W$  such that  $\not\models_w^{\mathcal{M}} \Phi$ . If  $\Phi$  is some propositional  $\varphi \in \mathfrak{Fml}$  or an effect law, any extension  $\mathcal{M}'$  of  $\mathcal{M}$  is such that  $\not\models_w^{\mathcal{M}'} \Phi$ . If  $\Phi$  is of the form  $\varphi \rightarrow \langle a \rangle \top$ , then  $\models_w^{\mathcal{M}} \varphi$  and  $R_a(w) = \emptyset$ . As any extension of  $\mathcal{M}$  is such that  $(u, v) \in R'$  if and only if  $u \in \text{val}(\mathcal{S}) \setminus W$ , only worlds other than those in  $W$  get a new leaving arrow. Thus  $(R \cup R')_a(w) = \emptyset$ , and then  $\not\models_w^{\mathcal{M}'} \Phi$ . ■

**Lemma B.4** *Let  $\mathcal{T}$  be modular,  $\Phi$  a law, and  $\mathcal{T}' \in \mathcal{T}_{\Phi}^-$ . If  $\mathcal{M}' = \langle \text{val}(\mathcal{S}'), R' \rangle$  is a model of  $\mathcal{T}'$ , then there is  $\mathcal{M} = \langle \text{val}(\mathcal{S}), R \rangle$  and  $\models^{\mathcal{M}} \mathcal{T}$  such that  $\mathcal{M}' \in \mathcal{M}'$  for some  $\mathcal{M}' \in \mathcal{M}_{\Phi}^-$ .*

**Proof:** Let  $\mathcal{M}' = \langle \text{val}(\mathcal{S}'), R' \rangle$  be such that  $\models^{\mathcal{M}'} \mathcal{T}'$ . If  $\models^{\mathcal{M}'} \mathcal{T}$ , the result follows. Let's suppose then  $\not\models^{\mathcal{M}'} \mathcal{T}$ . We analyze each case.

Let  $\Phi$  be of the form  $\varphi \rightarrow \langle a \rangle \top$ , for some  $\varphi \in \mathfrak{Fml}$ . Let  $\mathcal{M} = \langle \text{val}(\mathcal{S}), R \rangle$ . As  $\mathcal{T}$  is modular, by Lemmas B.2 and B.3,  $\mathcal{M}$  is non-empty and contains only models of  $\mathcal{T}$ .

Suppose  $\mathcal{M}'$  is not a minimal model of  $\mathcal{T}'$ , i.e., there is  $\mathcal{M}''$  such that  $\mathcal{M}'' \preceq_{\mathcal{M}} \mathcal{M}'$  for some  $\mathcal{M} \in \mathcal{M}$ . Then  $\mathcal{M}'$  and  $\mathcal{M}''$  differ only in the executability of  $a$  in a given  $\varphi$ -world, viz. a  $\pi \wedge \varphi_A$ -context, for some  $\pi \in IP(\mathcal{S} \wedge \varphi)$  and  $\varphi_A = \bigwedge_{p_i \in \overline{\text{atm}(\pi)}} p_i \wedge \bigwedge_{\substack{p_i \in \overline{\text{atm}(\pi)} \\ p_i \notin A}} \neg p_i$  such that  $A \subseteq \overline{\text{atm}(\pi)}$ .

Because  $\not\models^{\mathcal{M}'} (\pi \wedge \varphi_A) \rightarrow \langle a \rangle \top$ , we must have  $\models^{\mathcal{M}''} (\pi \wedge \varphi_A) \rightarrow \langle a \rangle \top$  and then  $\models^{\mathcal{M}''} \mathcal{T}$ . Hence  $\mathcal{M}'$  is minimal w.r.t.  $\preceq_{\mathcal{M}}$ .

When contracting executability laws,  $\mathcal{S}' = \mathcal{S}$ . Hence taking the right  $R$  and a minimal  $R_a^\varphi$  such that  $\mathcal{M} = \langle \text{val}(\mathcal{S}), R \rangle$  and  $R' = R \setminus R_a^\varphi$ , for some  $R_a^\varphi \subseteq \{(w, w') : \models_w^{\mathcal{M}} \varphi \text{ and } (w, w') \in R_a\}$ , we construct  $\mathcal{M}' = \mathcal{M} \cup \{\mathcal{M}'\} \in \mathcal{M}_{\varphi \rightarrow \langle a \rangle \top}^-$ .

Let  $\Phi$  be of the form  $\varphi \rightarrow [a]\psi$ , for  $\varphi, \psi \in \mathfrak{Fml}$ . Let  $\mathcal{M} = \langle \text{val}(\mathcal{S}), R \rangle$ . As  $\mathcal{T}$  is modular, by Lemmas B.2 and B.3,  $\mathcal{M}$  is non-empty and contains only models of  $\mathcal{T}$ .

We claim that  $\mathcal{M}'$  has only one arrow linking a  $\varphi$ -world, viz. a context  $\varphi_i \wedge \pi \wedge \varphi_A$  for some  $\pi \in IP(\mathcal{S} \wedge \varphi)$  and  $\varphi_A = \bigwedge_{\substack{p_i \in \overline{\text{atm}(\pi)} \\ p_i \in A}} p_i \wedge \bigwedge_{\substack{p_i \in \overline{\text{atm}(\pi)} \\ p_i \notin A}} \neg p_i$ ,

such that  $A \subseteq \overline{atm(\pi)}$ , to a  $\pi'$ -world, where  $\pi' \in IP(\mathcal{S} \wedge \neg\psi)$ . The proof is as follows: given  $\ell \in \mathfrak{Lit}$  such that  $\ell$  holds in this  $\varphi_i \wedge \pi \wedge \varphi_A$ -world

- if  $(\pi \wedge \varphi_A \wedge \ell) \rightarrow [a](\psi \vee \ell) \notin \mathcal{T}'$ , then  $\ell \notin \pi'$  and  $\mathcal{T} \models_{\overline{\text{PDL}}} (\pi \wedge \varphi_A \wedge \ell) \rightarrow [a]\neg\ell$ . Then this world has only  $\neg\ell$ -successors.
- if  $(\pi \wedge \varphi_A \wedge \ell) \rightarrow [a](\psi \vee \ell) \in \mathcal{T}'$ , then every  $\pi'$ -successor is an  $\ell$ -world.

By successively applying this reasoning to each  $\ell$  that holds in this  $\varphi_i \wedge \pi \wedge \varphi_A$ -world, we will end up with only one  $\pi'$ -successor.

Suppose now that  $\mathcal{M}'$  is not a minimal model of  $\mathcal{T}'$ , i.e., there is  $\mathcal{M}''$  such that  $\models^{\mathcal{M}''} \mathcal{T}'$  and  $\mathcal{M}'' \preceq_{\mathcal{M}} \mathcal{M}'$  for some  $\mathcal{M} \in \mathcal{M}$ . Then  $\mathcal{M}'$  and  $\mathcal{M}''$  differ only in the effects on that  $\varphi_i \wedge \pi \wedge \varphi_A$ -world:  $\mathcal{M}''$  has no arrow linking it to a  $\pi'$ -world. Then we have  $\models^{\mathcal{M}''} (\varphi_i \wedge \pi \wedge \varphi_A) \rightarrow [a]\psi_i$ , and then  $\models^{\mathcal{M}''} \mathcal{T}$ . Hence  $\mathcal{M}'$  is a minimal model of  $\mathcal{T}'$  w.r.t.  $\preceq_{\mathcal{M}}$ .

When contracting effect laws,  $\mathcal{S}' = \mathcal{S}$ . Thus taking the right  $R$  and a minimal  $R_a^{\varphi, \psi}$  such that  $\mathcal{M} = \langle val(\mathcal{S}), R \rangle$  and  $R' = R \cup R_a^{\varphi, \psi}$ , for some  $R_a^{\varphi, \psi} \subseteq \{(w, w') : \models_w^{\mathcal{M}} \varphi \text{ and } w' \in RelTarget(w, \varphi \rightarrow [a]\psi, \mathcal{M}, \mathcal{M})\}$ , we construct  $\mathcal{M}' = \mathcal{M} \cup \{\mathcal{M}'\} \in \mathcal{M}_{\varphi \rightarrow [a]\psi}^-$ .

Let now  $\Phi$  be  $\varphi$  for some  $\varphi \in \mathfrak{Fml}$ . Since  $\mathcal{T}$  is modular, by Lemmas B.2 and B.3 there is  $\mathcal{M} = \langle val(\mathcal{S}), R \rangle$  such that  $\models^{\mathcal{M}} \mathcal{T}$ . We know  $val(\mathcal{S}) \subseteq val(\mathcal{S}^-)$ . Because  $\neg\varphi \rightarrow [a]\perp \in \mathcal{T}'$ ,  $R'_a(v) = \emptyset$  for every  $\neg\varphi$ -world  $v$  added in  $\mathcal{M}'$ . Hence, because  $\ominus$  is minimal, taking  $\mathcal{M} = \{\mathcal{M}\}$  gives us the result. ■

### Proof of Theorem 5.3

From the hypothesis that  $\mathcal{T}$  is modular and Lemma B.1, it follows that  $\mathcal{T}'$  is modular, too. Then  $\mathcal{M}' = \langle val(\mathcal{S}'), R \rangle$  is a model of  $\mathcal{T}'$ , by Lemma B.3. From this and Lemma B.4 the result follows. ■